DESIGN OF AN INTERACTIVE VISUALIZATION SOLUTION FOR ANALYZING SPATIO-TEMPORAL CYCLING TRAJECTORY DATA: a case-study of the visualization of Fietstelweek data with BikePRINT

SHENGCE WANG
March, 2017

SUPERVISORS:
Dr. C.P.J.M. van Elzakker
Ir. M.J.G. Brussel

DESIGN OF AN INTERACTIVE VISUALIZATION SOLUTION FOR ANALYZING SPATIO-TEMPORAL CYCLING TRAJECTORY DATA: A case-study of the visualization of Fietstelweek data with BikePRINT
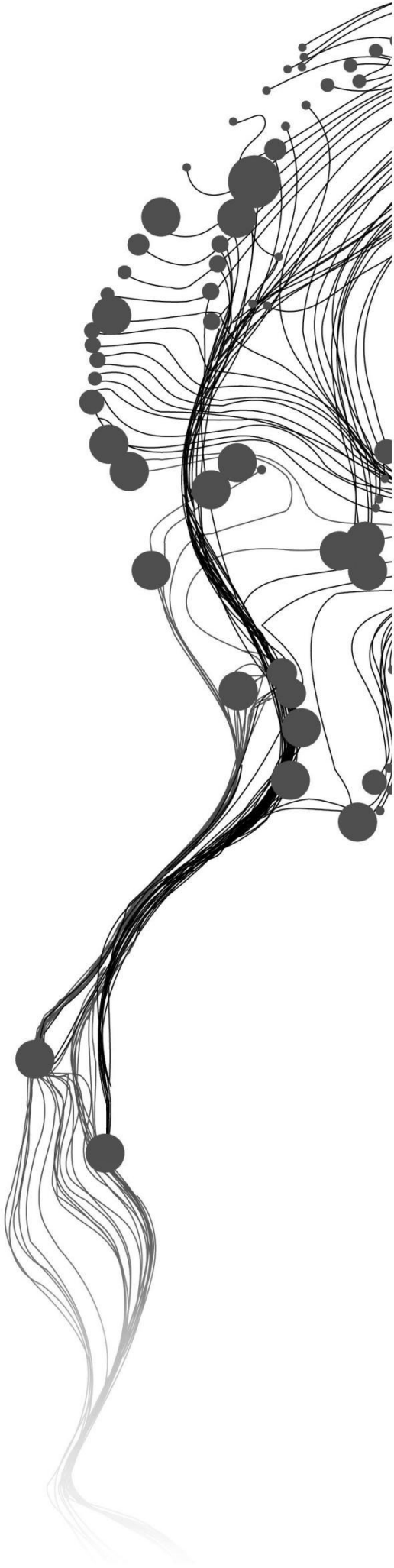
SHENGCE WANG

Enschede, The Netherlands, March, 2017

ABSTRACT

Analyzing spatio-temporal movement data is very popular and useful in urban studies. As one type of movement data, cycling trajectory data bring more opportunities for analysts to get insight into spatio-temporal information from different perspectives. Visualization is a useful tool which could be used by people to represent and explore the data, such as extracting spatial patterns. However, some existing visualization solutions have limitations that cannot assist people to do these works very well. So this research aims to find their limitations and design improved visualization solutions to help users exploring and analyzing data better. The stage of user research that investigates these limitations and acquires users' requirements with actual users is a very important part of the research. The recommendations of improved visualization solutions are made on the basis of the findings and results of this stage. In this work, several usability testing methods are applied to observe and evaluate how users use the existing visualization application BikePRINT to answer their analytical questions and complete their tasks. Firstly, questionnaire were sent to users to acquire their needs and interests. Then, 5 participants were invited to do analytical tasks under the recording. The results were analyzed and limitations of BikePRINT were Figured out. After that, improved visualization solutions were designed from three aspects according to the results of the user research: Accessing data, Visualization of accessed data and other functionality. These solutions were presented to and discussed with users and the developer in the focus group discussion. They gave comments, preferences and suggestions about these solutions. Accordingly, the final solutions were made sure.

Key words: Cycling trajectory, Movement data, Usability testing, User research, Visualization

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.   INTRODUCTION

1.1.   Motivation and problem statement

In order to decrease the environmental pollution and traffic congestion, in recent decades an increasing number of citizens chooses to ride bikes. With the development of advanced technologies like GPS trackers, wireless communication devices and other devices, recording the trajectories of cycling trips became feasible. In order to know cycling information better in the Netherlands, the 'Fietstelweek', one survey project of collecting cycling data, was initiated to collect these types of data as much as possible. These cycling trajectory data do not only represent the origin and destination of the cycling trips made by the cyclists but also the cycling trip's real trajectory. Hence, this provides new perspectives for a lot of analytical works related to cycling and cyclists on the basis of the understanding of movements of objects (Spaccapietra et al., 2008), such as analyzing the bike user's preferences of choosing paths (Beecham, 2015).

Cycling trajectory data describe how cyclists moved in cities which have been widely used by urban analysts/planners to get insight into spatial-temporal information of various areas such as urban traffic conditions, road using conditions, people's movements rules etc. Because of its spatio-temporal characteristic, visualization techniques can be used to represent trajectories directly which are commonly used to investigate the hidden information of bikes usage (Fisher, Dykes, & Wood, 1993), such as discovering bike users' behavioral characteristics, extracting spatial patterns (Jiménez et al., 2016). In the geography domain, we refer to visualization as GeoVisualization which refers to a set of tools and techniques supporting the analysis of geospatial data through the use of visualization tools (Wikipedia, 2016). With the help of visualizations, analysts can get insight into geographic data from more perspectives and even it is indispensable for some urban analysis (e.g. discovering the spatial pattern).
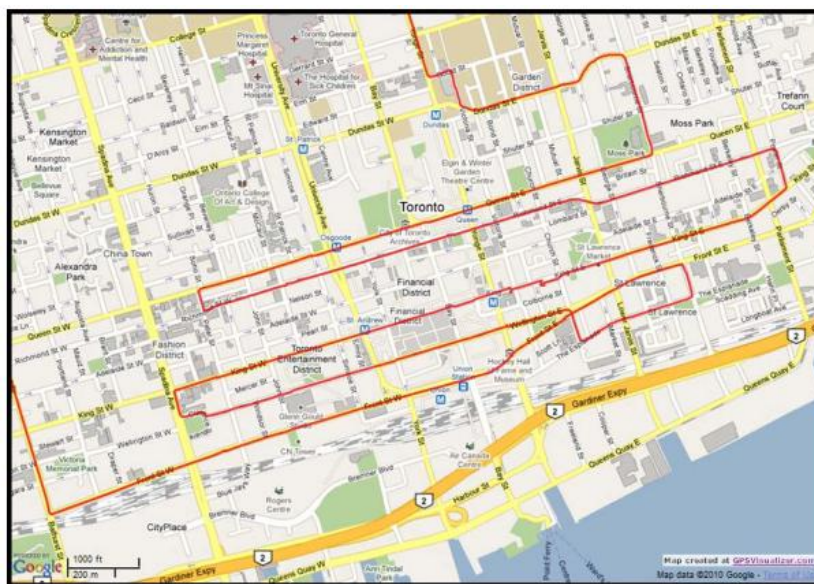
Figure 1-1 Red lines are drawn to represent a trajectory (Georgy & Noureldin, 2011)

There are many options to visualize trajectory data depending on the users' needs. Points, lines and their combination are always used to represent the trajectory. Taking Figure 1-1 as an example, red lines are

used to represent trajectories of movements of objects on a map. Visual variables like color, shape or size (thickness) of lines or points could be used to show other attributes and characteristics of the moving events and objects. On the basis of these representations, some visualization techniques were created by people to show the trajectory data. However, all these methods have their own advantages and disadvantages and people should have knowledge of this before they use them. When dealing complicated data or achieving complicated analytical tasks, scholars in geographic and cartographic domains are trying to integrate proper visualization tools and interactive methods into a multi-visualization application for supporting visual exploratory analysis of these spatio-temporal trajectory data (Andrienko, Andrienko, & Gatalsky, 2003). Such analysis refers to GeoVisual Analytics. Visual analytics can help users to represent spatio-temporal data by different points of view and manipulate them with several interactive methods. In these visual analytical systems, different visualizations are chosen to play different roles to meet users' different requirements. The application of interactive techniques makes the system responsive that brings more opportunities users to analyze the data more efficiently (Andrienko & Andrienko, 1999).

However, there are some challenges that remain when people use visualization methods to explore the cycling trajectory data. Different people have different purpose and tasks of analyzing the cycling data, or in other words, different people wants to extract different types of information from the data. The existence of these visualization methods of trajectory data leads that the people have difficulties to choose proper and suiTable visualization methods for their data or needs of analyzing data (Andrienko & Andrienko, 1999; Gonçalves, Afonso, & Martins, 2015). Choosing wrong methods leads they cannot visualize the data in a proper way, extract useful information effectively and the results may are not satisfied (Gonçalves, Afonso, & Martins, 2014a). In addition, time is indispensable for trajectory and people not only focus on comparing several patterns in different times or investigate changes in a specific time period but also are interested in temporal characteristics within one single trajectory. How to bring temporal information into spatial visualization is the main challenge (Boyandin et al., 2011). Moreover, Andrienko et al. (2003) pointed out that existing visualization tool commonly neglects the visualization of temporal characteristics of trajectory and the moving objects. Then, the insufficient usability testing of these systems after people produce them results in it may cannot meet target users' needs of analyzing the cycling trajectory data (Chen, 2005). Therefore, it is necessary to discover limitations of such existing visualizations and design improved interactive visualization solutions for exploring and analyzing these spatio-temporal trajectory data through using some user/usability testing methods towards the visualization.

1.2.   Research identification

1.2.1.   Research objectives
The main objective of this research is to design an interactive visualization solution for exploring and analyzing spatio-temporal cycling trajectory data.

There are 4 specific sub-objectives:
1)   To specify the characteristics of cycling trajectory data.
2)   To analyze existing visualization solutions for these data.
3)   To investigate the use and user requirements.
4)   To design and recommend improved visualization solutions.

1.2.2.   Research questions

Questions for sub_objective 1:

       Q1 What are spatio-temporal cycling trajectory data?

       Q2 What are the characteristics of these data?

Questions for sub_objective 2:

       Q3 What are the design solutions of visual representation of these data?

       Q4 What technical tools have been used in visualizing these data?

       Q5 What are the characteristics and (dis)advantages of these solutions?

Questions for sub_objective 3:

       Q6 Who are the users of spatio-temporal cycling data and for which purposes do they use the data?

       Q7 What analytical questions do they want to find answers to ?

       Q8 Which analytical questions cannot be answered well with the existing visualization solutions?

       Q9 What are the problems they meet when using the existing visualization solutions for their use cases?

       Q10 What factors lead to these problems?

Questions for sub_objective 4:

       Q11 According to the user requirements and existing solutions' limitations, what could be the design of improved solutions for the visualization of the data?

       Q12 What visualization tools can be used to design the visualization solution?

1.2.3.   Case study introduction

In this research, the focused data is 'Fietstelweek' (URL 1) data that provide the cycling information in the Netherlands. In that, information of cycling trajectories related to road network usage are contained in the dataset that could be used by people to analyze the traffic information or the cyclists' behaviors. Urban analysts are one of main user groups of this data and are also treated as the main users in this research. BikePRINT (URL 2) is a such visualization application made on the basis of 'Fietstelweek' data, which was made for translating the cycling data to spatio-temporal information to support urban analysts exploring and analyzing cycling trajectory data. However, it was told by the developer of BikePRINT that it remains some limitations that cannot help users to do analytical tasks very well and some of them are not willing to use it. So we would like to consider this application as the case study to research the visualization of cycling trajectory data and give improved visualization solutions.

1.3.   Thesis structure

In the thesis, Chapter 1 starts with briefly introducing the contents of the research including cycling trajectory data and its visualizations. Then, the challenges of visualizing this type of data are described as well and as a result, the objective and sub-objectives of this research that designing improved visualizations of cycling trajectory data are proposed. Accordingly, related research questions that lead this research are listed as well. In Chapter 2, the main purpose is to have a deeper understanding of related works of visualizing cycling trajectory data including: cycling analysis, trajectory data, current visualization methods of movement data, visual analytics and user research. From these works, the related works about visualizing trajectory data are known. For conducting the research well, some methodologies that could be used in the research should be made on the basis of literature. The planned methodologies used in the research will be introduced in Chapter 3. Then, In Chapter 4, the characteristics of 'Fietstelweek' data used

in this research are introduced first. There are some users who are using the data through its visualization application BikePRINT. However, BikePRINT cannot help users to answer their analytical questions or tasks very well. The description of visualizations of BikePRINT is made in this chapter. Then, in this chapter, an attempt is also made to list the different spatio-temporal analytical questions users of the Fietstelweek data may want to find answers to with the help of an application like BikePRINT. In order to understand the limitations of BikePRINT and acquire the users' needs of using 'Fietstelweek' data, user research with real users should be made. In Chapter 5, the main content is describing how to prepare and conduct the user research and the procedure of testing the limitations of BikePRINT is also described in detail, including research set-up, research implementation, research outcomes and analysis of the results. Then, according to the collected user requirements and limitations of BikePRINT, the conceptual design improved recommendations of visualization solutions will be made and described. In Chapter 6, improved possible visualization solutions will be given according the analysis of user requirements and the found limitations of BikePRINT. Finally, the conclusion of the research will be made and some possible future works are proposed in Chapter 7.

2.   LITERATURE REVIEW AND RELATED WORKS

2.1.   Introduction

This chapter introduces some basis theories and previous works related to this research. Firstly, the description of analytical works of cycling is made. Some works are based on the bike-sharing data which could also be supported by cycling trajectory data, such as extracting the spatial patterns and then analyzing cyclists' behaviors. While lacking some resource of cycling trajectory data leads people to have less understanding of what kind of analytical works could be done by using trajectory data. This research will investigate it and provide some findings of this. Then the characteristics of the cycling trajectory are introduced, including its concept, its representation, its main components and which of them could be visualized. Therefore, the ways of representing trajectory data visually are introduced in the next section. After that, current visualization methods of trajectory data are discussed that contain the static map, the animated map, the small-multiple and the space-time cube. Their advantages and disadvantages are also analyzed. In the next section, related visual analytics works of trajectory data are reviewed. As one of the important components, interactive techniques will also be introduced. Last part is about the related works about the user research.

2.2.   Cycling analysis

There are many types of research and analysis made on the basis of cycling data to get more insight of this green and popular transport mode. Wood, Slingsby, and Dykes (2011) used cycling data from the LCHS (London Cycle Hire Scheme) to identify changes in traveling behaviors with visualization approaches for improving the distribution of bike stations. Then, Beecham, Wood, and Bowerman (2012) used data from the same resource and developed useful tools to visualize and analyze user behaviors, such as exploring gendered cycling behaviors (Beecham & Wood, 2014b), identifying group-cycling behaviors (Beecham & Wood, 2014a), categorizing commuting behaviors (Beecham, Wood, & Bowerman, 2014) and improving cyclists' cycling experience (Beecham, 2015). Besides, Borgnat et al., (2009) visualized and explored the spatial pattern of cycling in Lyon to give a better understanding of the dynamics of the city and of human movements. Vogel, Greiser, and Mattfeld (2011) used Data Mining to acquire complicated cycling activity patterns to evaluate imbalances in the distribution of bikes. Then he also described cyclists' mobility and created user profiles of cyclists from extracting spatial patterns of cyclists to get a deeper understanding of conditions of cycling use in city Lyon (Vogel et al., 2014). The same works in Paris were made. Come et al., (2014) analyzed cycling and gave insights into the traveling behavior and the hidden urban phenomena linked to city dynamics, and Come, Randriamanamihaga, and Oukhellou (2014) extracted and visualized mobility patterns of cycling use in Paris. And Zhou (2015) analyzed the massive Bike-Sharing data in Chicago and offered an understanding of biking flow patterns by using data aggregation methods. However, a large part of these analytical works were made without considering trajectories of cycling because they used more bike-sharing data that only provide information of origins and destinations of cycling trips. Moreover, most researchers mainly focused on cycling behaviors but paid less attentions on cycling road usage, and connections among it, cycling behaviors and cycling policy. This is because the cycling data without trajectory information cannot provide the actual route of cyclists, so how do they behave during the cycling trip is unknown. And the relative lack of resource of cycling trajectory data is one of the reasons. The data used in the research brings opportunity to investigate these information. This

research could provide some understanding and findings about information mentioned above on the basis of case study dataset 'Fietstelweek'.

2.3. Cycling trajectory data

2.3.1. The concept of trajectory

Cycling trajectory data give people information of not only the origins and destinations but also the actual paths of cycling trips of cyclists which are a type of movement data that describes movements of moving objects among different locations, such as people, cars, ships, etc. The movement of the moving object describes how an entity changed its positions in a particular reference system (Andrienko et al., 2008). Generally, this reference system is a geographical system and we call the path of the entity's movement as the trajectory that is covered in a certain amount of time (Andrienko et al., 2008). Another explanation of trajectory data was made by Gonçalves et al. (2015), he defined that a spatial trajectory consists of a temporally ordered sequence of location points which is acquired from the registered spatio-temporal locations or extracted from other sources and some or all of moving points may contain some thematic data related to the moving events or the moving objects. Because the trajectory of movement cannot be established by the moving object in a flash, temporal components play an essential role in describing trajectories so that the trajectory can be considered as a series of spatio-temporal pairs (Andrienko et al., 2008). In the real world, time is continuous, so amounts of spatio-temporal pairs in a trajectory should be limitless, which, however, is not practical to be recorded by current positioning techniques and other sensors. Therefore, people record and represent trajectories by recording limited pairs to simulate the actual movements of moving objects.

2.3.2. Collection and pre-processing of trajectory data

As introduced in the previous section, people record several spatial-temporal pairs (points) to simulate the trajectory of the moving object by using current positioning techniques and other sensors. For example, mobile phones or devices equipped with GPS and wireless communication module could be used to record trajectories of humans, cyclists, cars, airplanes, etc. WiFi and Bluetooth devices could be used in indoor environments to record trajectories (Giannotti & Pedreschi, 2008). By these means, each moving object's trajectory is typically recorded and represented as several localized points of the tracked positioning device, called sampling (Giannotti & Pedreschi, 2008). According to Andrienko et al. (2008), there are five ways that trajectory data are recorded which make the sequence of locations of the trajectory:
1) People record the moving object's spatio-temporal pairs by regular time intervals or moments. For example: every 30 seconds or every 2 minutes.
2) People record the moving object's spatio-temporal pairs when it changes its location.
3) People record the moving object's spatio-temporal pairs when it moves into/close to a specific place. For example: when the observed object is close to the sensor, it records.
4) People record the moving object's spatio-temporal pairs when certain events occur.
5) People record the moving object's spatio-temporal pairs by combining these ways.

However, there are some disadvantages of these ways of recording and representing trajectories. The first is the measurement errors of current recording devices, especially in using GPS-equipped devices to make recordings or recording in GSM networks. Another disadvantage is related to the ways of recording points mentioned above in which the different sampling rates of different devices or ways of recording points affect the results of sampling. Reconstructing, clustering and classifying trajectories from these recorded point streams are needed (Greaves & Figliozzi, 2008). And then, some information could be extracted

from the trajectories, such as similar trajectories and some spatial patterns that reflect where moving objects frequently moved (Chen, Shen, & Zhou, 2011). For some moving objects moving on road networks (e.g. cars, cycling, trains and etc.), their trajectories also could be linked to these networks so that the usages of these networks are revealed. Or people may set many sensors on the locations of networks recording passing objects and their trajectories to meet these needs. In this way, on the one hand, the number of objects that passed the location could be recorded. On the other hand, which and how many segments of the network where a particular cyclist passed are known as well. Ahmed and Wenk (2012) introduced algorithms to construct trajectories and road networks from the unorganized point streams. Therefore, trajectories then could be aggregated and matched to segments of the network so that it could be represented by the sequence of segments where the moving object passed. Won et al., (2009) describe how they clustered trajectories in the road network and how they got high accuracy so that they could use a series of road names to represent the trajectory which makes readers know the trajectory better than using coordinates to describe it. In the 'Fietstelweek' data, the cycling trajectories are represented by a series of road segments, while the usage of road networks by cyclists is known as well.

The above different ways of recording brought various characteristics of the movement itself. The following sections will describe these characteristics in detail.

2.3.3. Characteristics of the trajectory

When dealing with trajectory data and visualizing it, the first thing is to find out what information could be represented in the trajectory. So, its characteristics are needed to touch with firstly. Andrienko et al. (2008) analyzed the trajectory data and proposed that investigating the characteristics of the trajectory could be from two aspects: (1) from the individual moment, (2) from movement in a period of time.

For the aspect (1): Characteristics of the former that are called moment-based characteristics contain, e.g. this information: the time, the moving object's location at this time, the direction of this moving objects' next movement, the instant speed of this movement, the change of direction and speed of the next movement and the accumulated time and distance of whole traveling at this movements (Andrienko et al., 2008). For the aspect (2): The another one is called overall characteristics of the trajectory which deals with a whole trajectory or a part of that has following information: geometric shape of the part of trajectory or the whole trajectory in the space, the traveling distance of the moving object, the costed time of the trajectory, the characteristics of the speed (the average/median/top speed, the periods when the speed increases/decreases/remains, the initial/final speed, the speed when the specific event happens, etc.), the characteristics of directions (the overall direction of the whole trajectory or the part of it, the initial/final direction , the direction when the specific event happens, the biggest angle of the change of the direction, etc.) (Andrienko et al., 2008).

The above descriptions of characteristics of the trajectory only focus on the single trajectory. However, dealing with two or several trajectories and finding their relations and connections are equally important. For example, works that from the level of several trajectories could be comparing trajectories of different moving objects (e.g. to discover the behaviors of two different taxis in a month), comparing the trajectories made at different times by one moving object (e.g. to compare trajectories of a person in two days) and comparing different parts of a whole trajectory (e.g. to compare trajectories of going to works and backing home of a cyclist). These relations could be the similarity and difference of trajectories, the order or the distance of trajectories and so on. In addition to this way, describing the trajectory could be

made from another perspective that are its three components: space, time and object (and its attributes) (Andrienko et al., 2011; Gonçalves et al., 2015).

2.3.4. The trajectory and space

In many fields of application, the moving object' position is more important than its dimensional entity so that moving objects are seen as moving points (Dodge, Weibel, & Lautenschütz, 2008). The spatial components of its trajectory are generally considered as series of points over time. Typically, trajectory data are always represented as a time-ordered sequence of location points $P =< x_n, y_n, z_n, t_n >$ where, x_n, y_n and z_n mean the geographic coordinates (latitude, longitude and altitude (if there is)) or three-dimensional positions in a non-geographic environment and t_n represents the corresponding times when the position was recorded. So a trajectory is represented by $T = \{P_1, P_2, ... P_n\}$.

In order to describe the trajectory, the space where the moving object is in should be determined at first. Space can be treated as 1, 2, 3-dimensional property according to the information to which the trajectory is related (Gonçalves et al., 2015). When only the distance between origin and destination of the trajectory is needed, it is enough to represent the space as a one-dimensional property. However, if it is important to show how the object moves over time, or to investigate relations among several different locations, 2D or 3D (with height information) representation are more suiTable.

2.3.5. The trajectory and time

Time is another critical component in describing cycling trajectory data which helps us to investigate the data from the past to the future (Aigner et al., 2007). It is impossible that moving objects can make trajectory instantly, so that time is an indispensable element of a movement (Andrienko et al., 2008; Miller, 2005). Representing time needs consideration of its hierarchical system of granularities, consisting of seconds, minutes, hours, days, weeks, months, years, decades, centuries, and so on (Andrienko et al., 2010). Time contains (1) the natural cycles in that is relatively regular and could be predicted to a certain degree, such as the season; and (2) the less regular re-occurrences, such as the social, economic or natural cycles (Gonçalves et al., 2015).

There are two main aspects of the temporal dimension that should be considered when dealing with spatio-temporal or time-oriented data (Frank, 1998). The first one is the time primitives that constitute the temporal dimension, which is categorized into time points (an instant of time) or time intervals (a period of time), while second is the organizational structure of the temporal dimension where there are three types: ordered time, branching time, and multiple perspectives (Gonçalves et al., 2015). For describing the temporal component of the trajectory/trajectories, ordered time is considered mostly so we explain this here. The ordered time still can be subdivided into two types: linear and cyclic time, depending on whether the time is considered as an ordered series of temporal primitives or a set of recurring temporal primitives. Generally, the trajectory is made following the linear series of time primitives, while the cyclic time is related to moving objects' activities from the overall respective, for example, cyclists' movements occur in the daily cycle (Andrienko et al., 2008).

2.3.6. The moving object and its attributes

No matter where these trajectory data were collected, with current techniques we are also able to record moving objects' other attributes than just spatial and temporal information, such has the object type, the object size, or the altitude (Scheepens et al., 2011). Attribute information is commonly non-spatial, non-

temporal and links to specific themes (the daily trips of employees, the migrations of birds and the trajectories of bikes), generally, which could be divided into two types: quantitative (e.g. the instant speeds in different times) and qualitative (e.g. the type of cyclists) (Qian et al., 1997). Also, some characteristics of the movement made by moving objects, such as speeds, directions and accelerations should be concerned (Giannotti & Pedreschi, 2007). Therefore, it is also important to treat these attribute information of the moving object(s), movements and its environment well because a better visualization of this information contributes to a better analysis of a spatio-temporal phenomenon of the trajectory (Roth, 2012).

For example, the moving behaviors of cyclists may depend on the characteristics of the cyclists (the gender, his/her job, age, health conditions, etc.), the environment when he/she rides bike in (the road quality, the traffic control, the climate, etc.), the characteristics of his/her bike (the type, the condition, etc.) and so on, which all will affect the characteristics of the trajectory (Andrienko et al., 2008). According to Stacking-based visualization of trajectory attribute data, it is not easy to represent and understand these attributes that are related to spatio-temporal trajectories because it involves many complex elements. For example, the attributes should be analyzed not only at the level of the single trajectory, but also at the level of the group of several trajectories. In general, whether the attributes of the moving object and the trajectory can be expressed well or not will affect how much true meanings of the moving object and the movement could be represented (Tominski et al., 2012).

2.3.7. Conclusion

The trajectory and its characteristics could be described from many points of view, from the interior of the trajectory, the whole trajectory and to the group of trajectories, or from the components of the trajectory and the moving objects that create it. Accordingly, people's needs of investigating and exploring these characteristics vary. Therefore, different visualizations could be chosen to meet these requirements.

2.4. Visual representations of trajectory data

2.4.1. Introduction

Since the raw format of data cannot provide users with much meaningful information, some adequate ways should be used to represent data by people to get insights from the data. Because visual representation of data could increase people's perception and cognition to the information, data visualization became an important approach to support people exploring and analyzing data (Andrienko et al., 2008). Due to the spatial characteristics of cycling trajectory data introduced above, maps are often seen as main and important tools for visualizing trajectory data (Kraak, 2008). And in visualizing trajectory data, the map is the main or one of the main component(s) in the whole visualization solution, while the other non-map visualization methods play additional roles to represent the attributes of trajectory (Gonçalves et al., 2015). In following sections, the ways of representing the trajectory on the map will be introduced.

2.4.2. Trajectory data and OD data

For visualizing origin-destination (OD) data, people only need to consider about two locations of the movement. However, when representing a trajectory, not only the origin and destination of the trajectory should be visualized, other points and connected segments that constitute this trajectory are also could to be represented to describe the trajectory. Figure 2-1 is a simple illustration of showing the difference of the spatial representation of trajectory data and OD data. As we can see, the red line is used to represent OD data. Only the origin and destination are considered, the procedure of the moving is ignored. And

blue lines are used to represent the trajectory from point 1 to point N. These two representations only are the spatial representations that don't consider other attributes of the movement. However, there are many attribute data related to the movement, the moving object and the time that could also be represented (e.g. instant speed, direction, types of objects). Visual variables can be used to represent these data (Figure 2-2). For example, the size could be used to represent the number of objects passing by a certain road segment, the color hue could be used hue to represent the nature of a moving object and the color value could be used to show average speed.
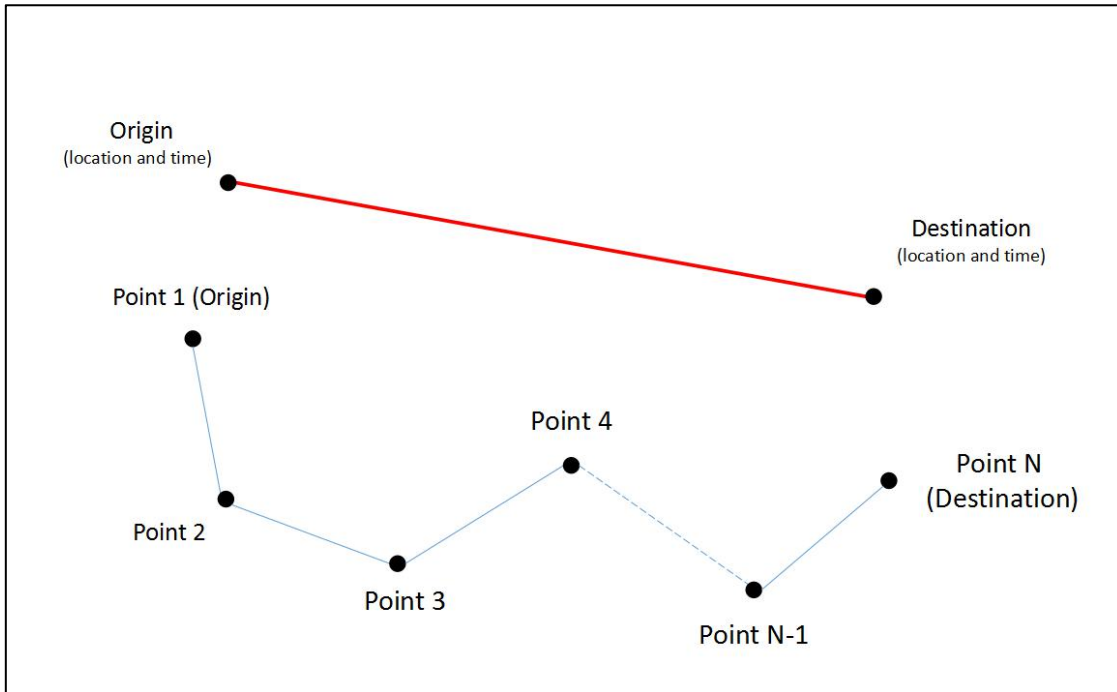
Figure 2-1 Difference between visual representations of OD data and trajectory data

Figure 2-2 Visual variables (Bertin, 1983)

2.4.3. Ways of representing the trajectory

Firstly, the way of representing spatial components of the trajectory will be discussed. As the trajectory is mainly reconstructed and extracted by recording spatio-temporal points, it could be directly represented by these points if we do not consider the shape of the moving object. Figure 2-3 is a simple illustration of representing the trajectory without representing attributes of the movements. As we can see, every point represents one location where the moving object passed by which could be the geographic coordinate or the location of a space. Interpolating points could increase the density of points that describe the trajectory smoother (Figure 2-4) (Macedo et al., 2008). Attributes of movements can be represented by applying different visual variables to these points. For example, different sizes of points can represent the instant speed at this point. However, these points cannot represent the trajectory continuously if there aren't high sampling rate of points.
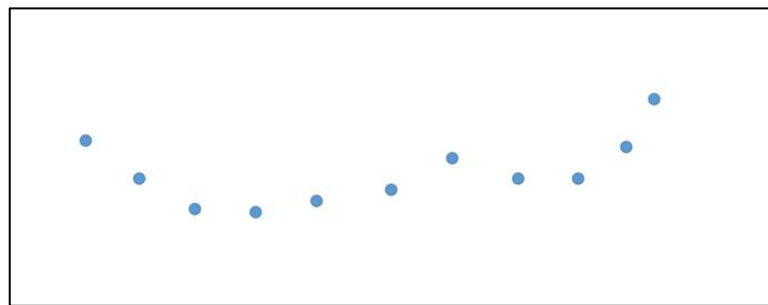
Figure 2-3 Using recorded points to represent the trajectory
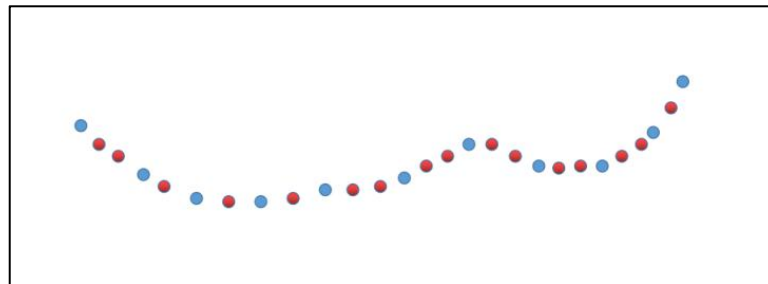
Figure 2-4 Using points to represent the trajectory after interpolation in which red points are interpolated points.

Using lines to connect these points is another way of interpolation, which makes the trajectory to become continuous. Lines could be straight lines or curved lines called linear interpolation or Interpolation with Bezier curves (Macedo et al., 2008). Then, drawing points and lines on the map together to represent the continuous view of the trajectory. This way for visualization is similar to using the flow map to visualizing OD data. The first point is seen as the first origin point, while the second point can be seen as the first destination point and also the second origin point. The rest can be done in the same manner. Figure 2-5 is a simple illustration of representing the trajectory by points and interpolated lines. The attribute of the movement can be represented by visual variables applied to points and lines. For example, the average speed between two recorded points could be represented by the visual variable lightness value or with the visual variable width of the interpolated line.
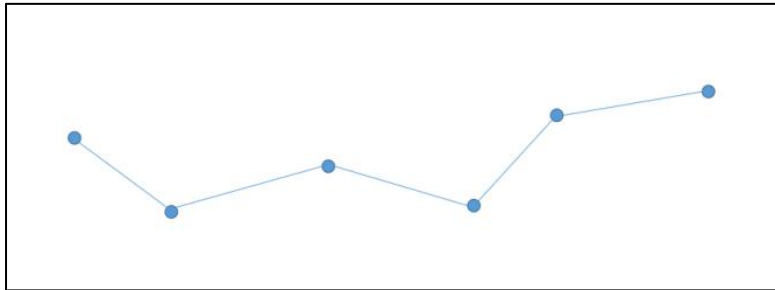
Figure 2-5 Using lines and points to represent the trajectory

Sometimes, the display of points can be omitted when the information of points is not important or unconcerned. Only using lines to represent the trajectory is another way. Figure 2-6 is a simple illustration of this way of representing the trajectory. For those trajectories which are aggregated and matched to the road network, this way could be used to represent the trajectory, which is used in 'Fietstelweek' to represent cycling trajectories. Lines for representing trajectories are a sequence of road segments. In Figure 2-7, the highlighted blue road segments (lines) are together building a cycling trajectory in the network with a red color. In this representation, the visual variables of these road segments (lines) could be used to show not only attributes of the movement but also attributes of the road segments which are related to the movements (e.g. the amount of cyclists on the road segment as shown by Figure 2-8).
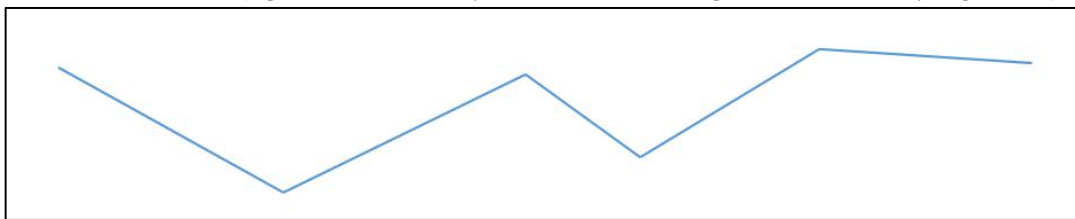
Figure 2-6 Only using lines to represent the trajectory
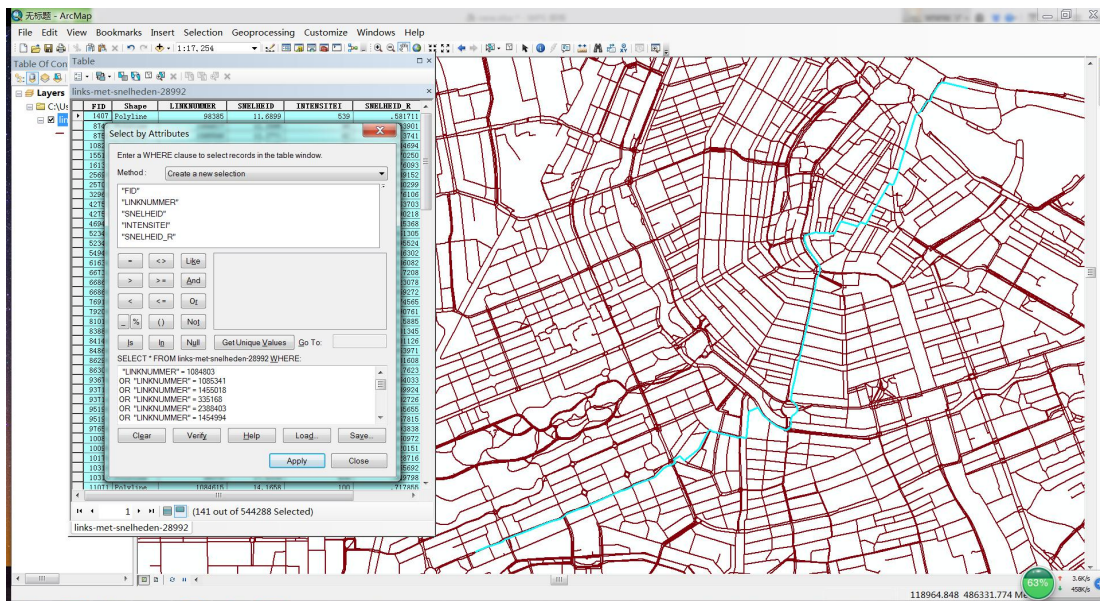
Figure 2-7 A sequence of highlighted blue lines are used to show a cycling trajectory in the road network
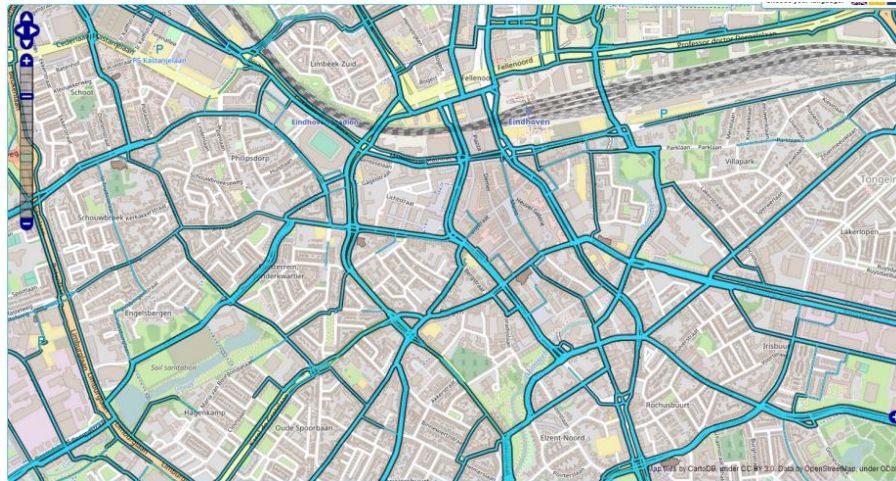
Figure 2-8 The width of road segments represent different amounts of cyclists

Over past decades, many visualization methods were created to represent trajectory data on the basis of these visual representations (Andrienko et al., 2013). Every visualization method has its pros and cons that depend on specific data types and tasks so that before choosing and applying visual representations, it is essential to understand their pros and cons (Kraak & Madzudzo, 2007). Gonçalves et al. (2015) classified these visualization methods to four types depending on how the temporal information of trajectory are dealt with: (1) the static map, (2) small-multiple maps and animated maps and (3) the space-time cube, which will be described in following three sections.

2.5. Visualization methods of trajectory data

2.5.1. The static map

Using static maps is one of the earliest ways of visualizing data of a specific topic for several decades in which the visual variables are used to represent changes in different times, represent an event that occurs at a certain time moment or show the properties of objects in the map (Biadgilgn, Blok, & Huisman, 2011). It was drawn by hand firstly (Phan et al., 2005) and then Tobler (1987) gave people some instances of how the computer can be helpful to produce the maps for visualizing migration data. As described above, the trajectory could be drawn as two-dimensional lines on the static map. Or if the data were aggregated to some network data, highlighting the network segments which the moving object passed is another way to represent the trajectory. The changes of visual variables, such as colors, shapes, widths, transparency and some symbols are used to show changes of attributes of moving objects or the movement (Kraak & Ormeling, 2011). With the help of geographic information technology, people could represent temporal information of the movement on the static map. However, according to (Gonçalves, Afonso, & Martins, 2014b), for representing temporal components, some additional visual elements (e.g. a timeline) are needed to combine with the static map. And Gonçalves et al. (2015) also indicated that temporal components may are most easily unnoticed by using the static map. Figure 2-9 is a static map for visualizing the trajectory.
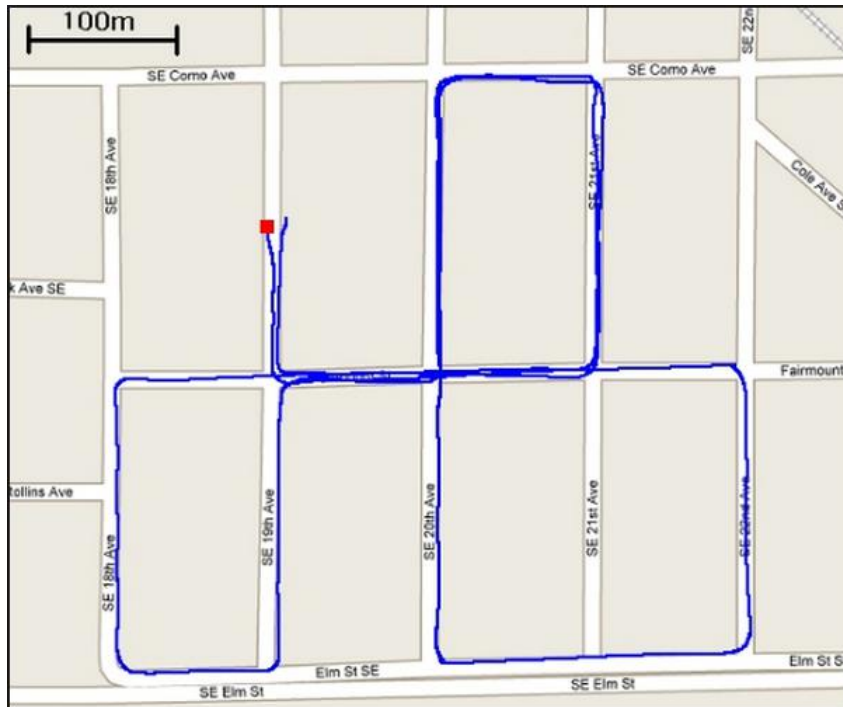
Figure 2-9 A static map of visualizing the trajectory (URL 3)

However, because of its nature, it is relatively good for using static maps to deal with one or a few trajectories of movements (Goesele et al., 2012), but when the amount of data increases largely, the overlaps of massive trajectories may lead to a cluttered display (Figure 2-10) which affects the users' acquisition and understanding of spatial information (Elmqvist & Fekete, 2010; Zhou, 2015). Using a spatial generalization method is one of useful approaches to avoid cluttered displays (Guo, 2009; Phan et al., 2005), but accordingly will lead some loss of information (Guo & Zhu, 2014). Another one is using data selection (filtering) by some parameters to acquire a little part of data (Andrienko et al., 2011). Therefore, only using a static map to visualize trajectories is not qualified. The combination of the static map and other data visualization methods is recommended (e.g. parallel coordinates, time graph and etc.). Moreover, other methods that developed from the static map are alternatives to visualize trajectory data which will be described below.
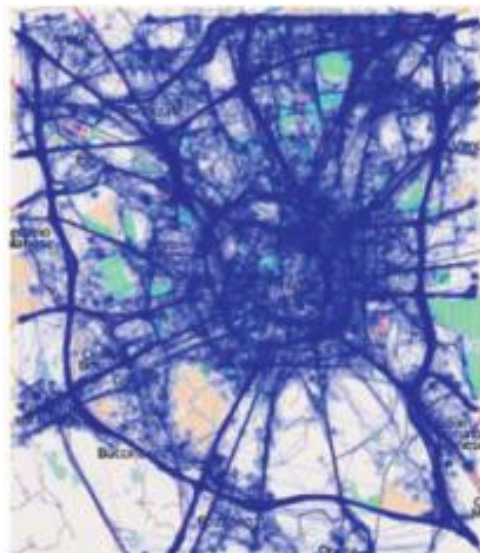
Figure 2-10 Cluttered display of trajectories (Adrienko & Adrienko, 2011)

2.5.2. Small-multiples and Animated maps

With the help of the rapid development of GIS (geographic information system) technology and computer science, many new methods were developed to help people to visualize and explore this type of data on the basis of the static map. There are two techniques are used commonly: Small-multiples and Animated maps. The purpose of using these two techniques are same: using several snapshots to show different attributes or processes of the trajectory which focus on emphasizing changes and the comparison (Gonçalves et al., 2015). So we present them together.

As shown by Figure 2-11, using small-multiples to visualize data is a way of combining several static maps that respectively represent different information in different certain moments (Kraak & Ormeling, 2011). In this Figure, every small map could represent a moving state of objects in a trajectory in different times. In this way, people investigate the change of spatial data by following the time sequence (Gonçalves et al., 2013). However, Boyandin (2013) and Ostermann (2010) pointed that if the amount of moments/small maps increases, the size of each small map deceases. This will make people looking at the information in maps with more difficulties size that some detailed information may lose.
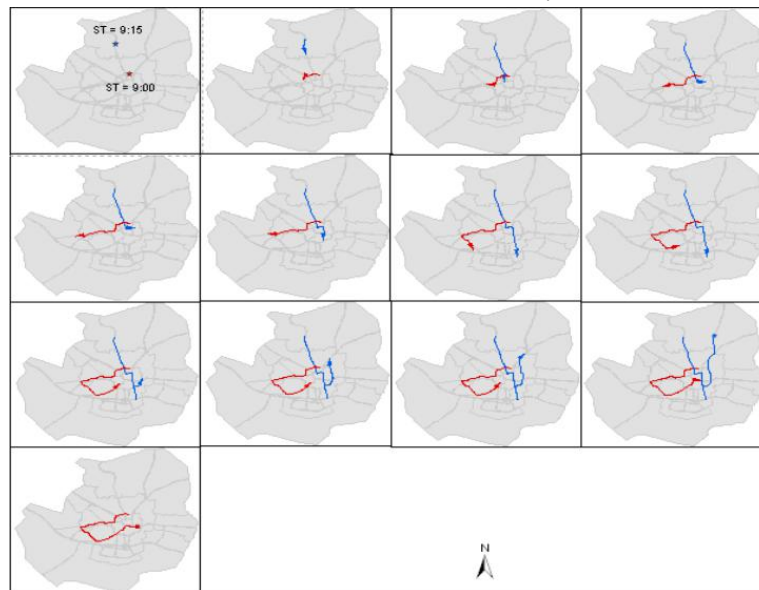
Figure 2-11 Small multiples of showing changes of trajectories (Biadgilgn et al., 2011)

The animation map represents the change of the movement by displaying a sequence of images one by one in a single view (Boyandin, 2013). According to Kraak and Ormeling (2011), there are two main types of animated maps: time-based or non-time-based. For example, time-based animated maps could be used to show the change of moving state of objects in a trajectory in different times (Figure 2-12), while the non-time-based animated maps could be used to compare the patterns of different types of trajectories by placing them next to each other. Besides, the animated map could play automatically for people to have an overall understanding of the change and the user also could control it to switch to the moment they want for having a detailed look of information (Boyandin, 2013). However, using the animated map is less effectively and time consuming if there are big amount of moments needed to be dealt with.

Figure 2-12 The animated map used to show how an object moves (Gonçalves et al., 2015)

2.5.3.   Space-Time Cube

Space-Time Cube (STC) is a developed visualization method that uses the three-dimensional visualization to represent both spatial and temporal information of movements within a cube, where the x and y axis represent the geographic locations and the third dimension z-axis is used to represent time (Gonçalves et al., 2014b; Kraak, 2003b). According to Kveladze, Kraak, and Van Elzakker (2015), the Space-Time Cube could help people to represent the space, time and their relationships of the movement at the same time. It is also seen as a special type of three-dimension visualization because the third dimension represent the temporal information (Kjellin et al., 2010a). However, because people' s perceptual issues towards to 3D properties of the space-time cube, when dealing with massive trajectories, some clutters may be generated that leads to the undesirable visualization result (Gonçalves et al., 2014b; Kjellin et al., 2010a). Therefore, interactive methods are always combined with the space-time cube for users to filter the data or manipulate the map, change the point of view of observing the visualization (Kjellin et al., 2010b). The Space-Time Cube was originally developed from the space-time model introduced by Hägerstrand which is seen as the the starting of time-geography studies. Time-geography is a study about about people's movements with respect to space and time.

The space-time cube is always used to describe movements with respect to the temporal information (Kveladze, 2015). Some visualizations of movements have been made by the help of the space-time cube. For example, as Figure 2-13 shows, Kraak (2003a) produced the visualization of the classic static map mentioned above - Napoleon's campaign to Moscow in the space-time cube environment. Kwan and Lee (2004) visualized the patterns of human movements but using the space-time cube. He also concluded that this method not only can extract spatio-temporal information of human's moving behaviors, but also it is helpful for modeling these behaviors. Andrienko et al., (2014) visualized trajectories of cars in the city by using the space-time cube technique in which his focus was attributes of these trajectories. As the result, the original space-time cube has been changed to a new form called 'trajectory wall' (Figure 2-14).
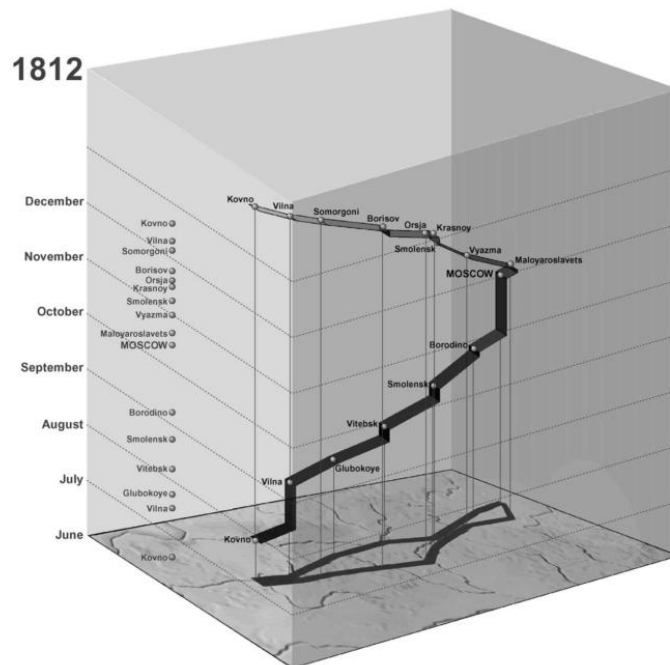
Figure 2-13 Using space-time cube to show Napoleon's campaign to Moscow (Kraak, 2003a)
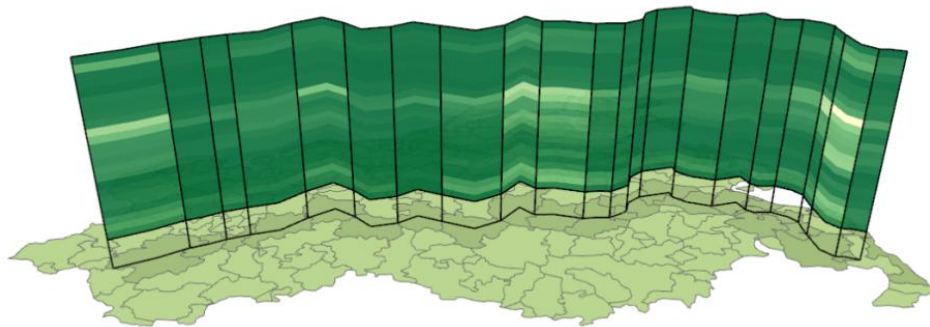
Figure 2-14 The 'Great Wall' to show the trajectory, the wall could be used to show attributes (Goesele et al., 2012)

The space-time cube could provide a three-dimensional environment for users to represent spatio-temporal movement data, especially the representation of temporal component is its strong point. The single trajectory is clear shown in the environment as a path. However, if the data size increases, due to its own properties, it likely leads to the visual clutters. Therefore, interactive methods are always integrated when producing space-time cube environment. In this way, users could manipulate the map like choosing datasets, comparing or filtering the data in the space-time cube environment.

2.5.4.   Summary

These sections introduced current visualization techniques which could be applied to visualize trajectory data. Their advantages and disadvantage should be considered when people choose them to visualize the data. The attributes of the trajectory also could be visualized by the combination of these methods. The direction of the movement could be represented in the trajectory. The direction describes how moving objects moved and how the movement developed. Here, we firstly discuss how to represent trajectory's directions statically on the single map, and then describe how to represent directions with the help of other visualization techniques.

2.6. Visualizations of the direction of the movements

Directions are key information of movements that describe how moving objects moved and created trajectories. In addition, when urban planners do the analysis at the level of several trajectories, their directions at different moments may reflect the trend of movements. With the help of dynamic visualization techniques, it is relatively easy to represent directions of the trajectory. When representing a trajectory in a single static display, arrows and particular visual variables are two main ways of representing the direction.

2.6.1. Static representation of the direction of the movements

Points and lines are two important components that mainly constitute the trajectory. Therefore, we describe how to use arrows and visual variables to represent directions from these two types.

1) For points, using visual variables (different shapes, sizes and colors) of points is the main choice of representing the direction. Using the arrow as the shape of the point is also possible (Figure 2-15). In Figure 16, there are some examples of using visual variables.
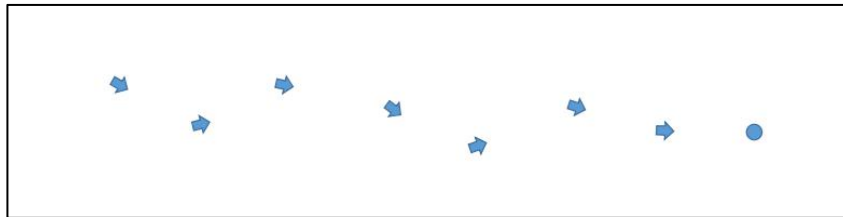
Figure 2-15 Using arrows as the shapes of points
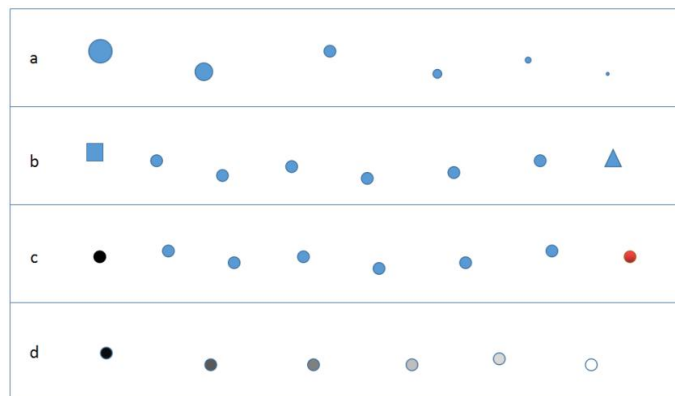
Figure 2-16 Some examples of using visual variables to represent the direction

- Figure 2-16(a) is using different sizes of points to show the direction of trajectory. The biggest point is the origin and the smallest is the destination. The direction of the movement is from the bigger point to the smaller point.
- Figure 2-16(b) is using different shapes of origin and destination points to show the direction of a trajectory. The square can be seen as the origin point and the triangle can be seen as the destination point. So the direction of the trajectory is from the square to the triangle along other points.
- Figure 2-16(c) is using different color hues of origin and destination points to show the direction of trajectory. The black point means the origin and red one is the destination.
- Figure 2-16(d) is using different color lightness values of points to show the direction. The direction could be from the dark point to the light point.

These ways truly can represent the trajectory by some ordered value or specific shapes or colors to define the origin and destination of the movement. However, the results of visual representation are not very desirable. In addition, if there are many trajectories on the map simultaneously and even they could overlap each other, these representation ways are worse because readers cannot distinguish them well.

2) When using lines to represent the direction, visual variables and arrows of lines can be applied to represent the direction as well. Arrows can be in the front of lines, in the middle of lines or inside of lines. In Figure 2-17, there are examples of these three types of arrow representation. The points in the Figure have no contribution to representing the direction.
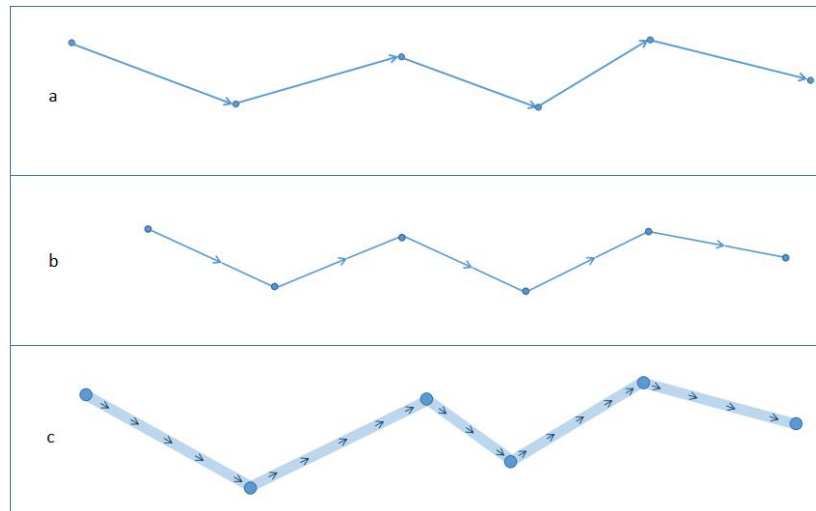
Figure 2-17 Three types of using arrows to represent the direction

- Figure 2-17(a) is putting arrows at the end of each line segment.
- Figure 2-17(b) is putting arrows in the middle of line segments.
- Figure 2-17(c) is embedding arrows into line segments.

The directions that arrows point at are the directions that the moving object moved to.

Using visual variables of lines is another way to represent the direction. There are some examples in Figure 2-18.
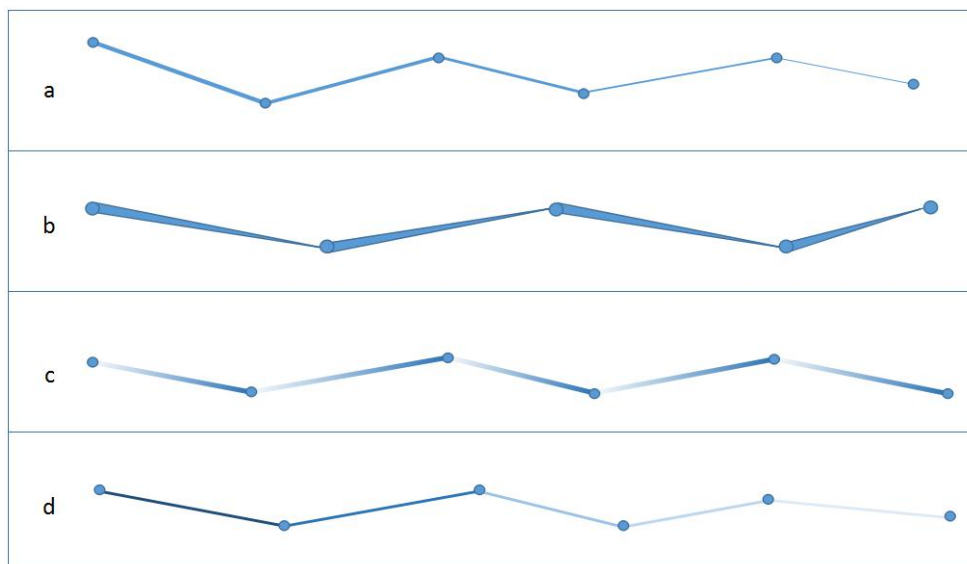
Figure 2-18 4 examples of using visual variables applied to line symbols to represent the direction

-      In Figure 2-17(a), different widths of all lines represent the direction that could be from the wide line to the thin line.
-      In Figure 2-17(b), different widths vary in a single line that represent the direction from the wide side to thin side.
-      In Figure 2-17(c), different lightness values in a single line represent the direction.
-      In Figure 2-17(d), different lightness values of all lines represent the direction.

2.6.2. Representing direction with other visualization techniques

The direction also could be represented by applying the small multiples, animated map, and space-time cube to represent the movement.

1) Small multiples and animated maps

Small multiples and animated maps can represent the trajectory by steps of movement with respect to the time. The direction of movement can be extracted by comparing these small maps. Lines and points both could be represented in this way. Since the contents of every map of animated maps are the same as the small multiples, we only demonstrate the display of small multiples. Figure 2-19 (lines) and Figure 2-20 (points) are two simple illustrations of representing the changes of trajectory made by the moving object at different time moments. From this, we can know the direction of the trajectory.
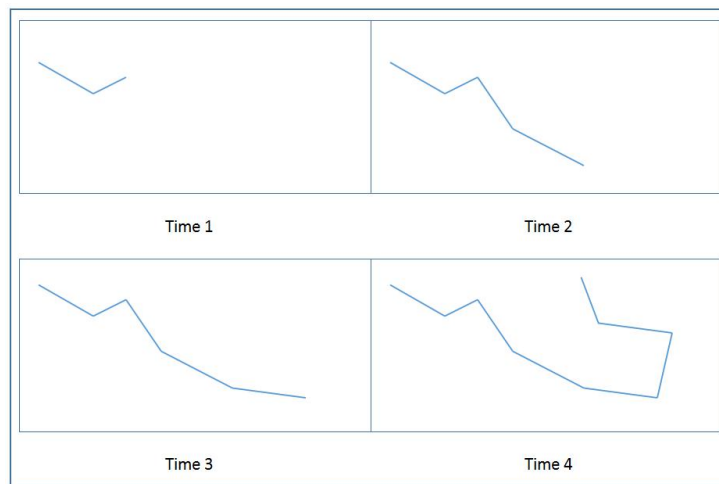
Figure 2-19 Representing the object's movement by lines in small multiples
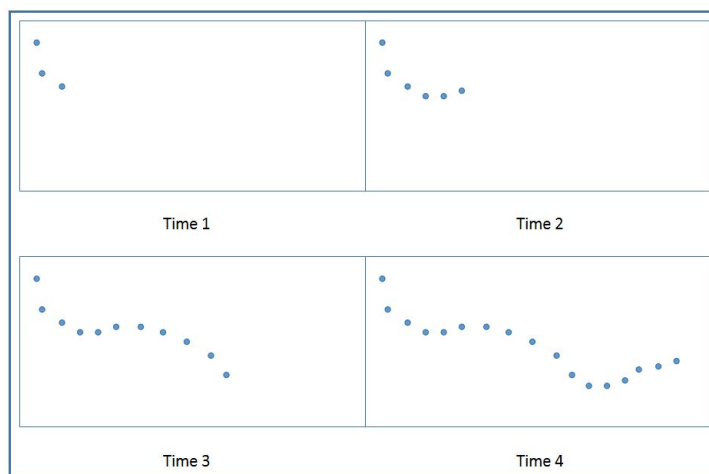
Figure 2-20 Representing the object's movement by points in small multiples

In some cases, trajectories are accumulated or the road network are known beforehand. Then, locations of the moving object at different time moments are shown in different maps. As exemplified in Figure 21, the blue lines represent the whole trajectory or the road network, while the black point means the moving object.
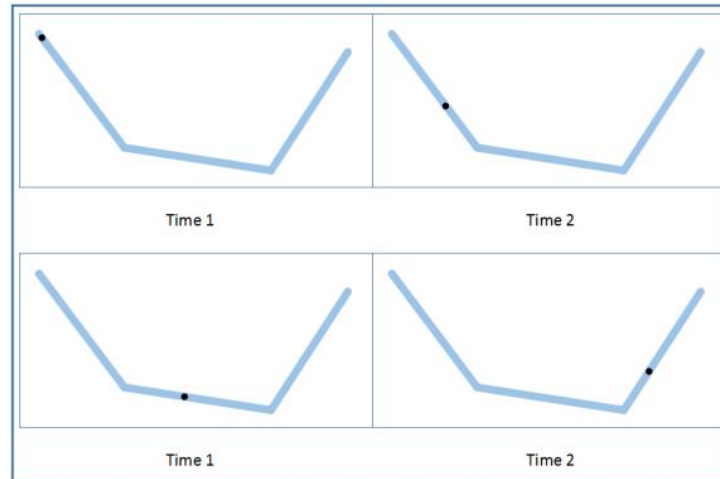
Figure 2-21 Representing the object's movement by different locations of the moving object on the network in small multiples

2)   Space-Time Cube (STC)

The methods described above are both 2D representations. In a 3D environment, the dimensions X and Y of a STC could represent the spatial components of the trajectory and the vertical dimension Z represents time. When the trajectory is visualized, its direction could also be extracted from the change of time in the Z dimension. However, if the third dimension is used to represent the actual height of the trajectory, arrows and visual variables of the trajectory (lines and points in the 3D environment) are needed to represent the direction.

2.6.3.   Conclusion

In this introduction, we introduced the representation of the trajectory. When representing the trajectory by a single static display, visual variables and arrows are two main alternatives. However, the visual variables of points and lines are more used to represent other quantitative and qualitative attributes of the movement, using arrows is a recommended way to represent the direction. Beside to the static map, other visualization techniques, such as small multiples, animated maps and Space-Time Cube also could be used to visualize trajectory data.

2.7.   Visual Analytics and Interaction

This section introduces the concept of visual analytics and how it helps people to explore the spatio-temporal data. In visualizing and analyzing trajectory data, some works could be done with the help of visual analytics are also described. Then the essential component of visual analytics, the interaction, is introduced because evaluating the interactions in the visualization environments is also important. Finally, some previous visual analytics works of trajectory data are listed. How they help people exploring and visualizing trajectory data are also described.

2.7.1.   Visual Analytics of the trajectory

There is a challenge that pointed out by Andrienko et al. (2008) that using some visualization methods to only help people investigating the trajectory data from narrow perspective. However, from the movement

itself, to the characteristics of moving objects and the environment where the moving object moves in all influence the characteristics of trajectory data. Moreover, in order to deal with increasingly complicated user tasks, it is necessary to choose suiTable visualization methods and combine them with the consideration of several aspects including the backgrounds and characteristics of users, their needs, tasks and so on that could effectively support people's analytical works of movement data (Andrienko & Andrienko, 2007). The visual analytics is such a combination that facilitates human's cognitive abilities of information visualization and computers' abilities of analysis, which helps people to understand the trajectory data efficiently (Ren et al., 2014). In order to better analyze movements by the help of visual analytics, N. Andrienko and Andrienko (2013) proposed that these analytical works towards movements could be grouped to four types: (1) Focusing on the movements by analyzing characteristics within a trajectory. In this way, the trajectory is see as the combination of several segments and points. The characteristics of moving activities are important components investigated in this way; (2) Considering the trajectory as a whole. This type of analysis is intended to explore the spatio-temporal changes of a single trajectory or the comparison of these information of several movements; (3) Focusing on the overall spatio-temporal distribution made by several movements. The single trajectory is not the emphasis in this kind of work. (4) considering the movement(s), the environment where the moving objects move and the relations and connections between them.

2.7.2. Interactive technology

Interactive technology is a bridge between the visualization approach and the user that plays a very essential role in the visual analytics (Yi, ah Kang, & Stasko, 2007). Through interactive methods, users could explore the data by their interests or needs. The concept of human-computer interaction (HCI) is the core in understanding interactive methods. As the introduction made by Dix (2009), the purpose of the interactive system, the expert domain, the user tasks, goals, analysis of tasks, user's and system's language and the system itself are all need to be considered when designing and dealing with the interaction. In addition, having a better understanding of taxonomies of interactive methods makes people could design the interactive methods in visualization system for meets user's needs of exploring data and doing the analytical tasks effectively (Yi et al., 2007). One of the most classic taxonomies of interaction was made by Shneiderman (1996), which are Overview, Zoom Filter, Details-on-demand, Relate, History and Extract. Beside to this way of classifying interactive methods, other researchers proposed their ideas of taxonomies of interaction. Some of them made the taxonomy from the low-level (Buja, Cook, & Swayne, 1996; Chuah & Roth, 1996), while some classified from the view of dimensions Information (Spencer, 2007; Tweedie, 1997) and others considered the interactive methods following taxonomy of user tasks (Amar, Eagan, & Stasko, 2005). However, these classifications have many similarities and occasionally some different interactive methods are used to deal with the same user tasks or purposes. So, Yi et al. (2007) proposed that task-oriented taxonomies of interactive methods are more effectively to deal with and on the basis of this theory he also grouped the interactive methods in visualizations to seven types: 1) Select, 2) Explore, 3) ReconFigure, 4) Encode, 5) Abstract/Elaborate, 6) Filter, and 7) Connect through the analysis of several visualization systems and get insight into the characteristics of interactions.

2.7.3. Existing visual analytics systems for trajectory data

Analytical works of movement data are very popular in visual analytics and various visual analytics methods and tools have been created for analyzing and exploring the movements (Andrienko & Andrienko, 2013). By these means, people explore the data from different perspectives and answer analytical questions, especially when the amount and complexity of movements data increase. The visual

analytics is applied in various types of data and domains. However, in this research, we mainly focus on some existing visual analytical system with respect to trajectory data.

Vistradas was made by Barbosa et al., (2014) which is a visual analytics system that helps people to understand trajectories data of public buses in the city of Rio de Janeiro so that the related authorities can better monitor the bus services in the city. In the works, they visualized the data of expected bus routes and GPS data of actual bus routes on the map to make the verification by comparing them. In addition, they also analyzed testing sameness of several buses and investigated how events influence buses' movements.

Gonçalves et al. (2013) introduced how they created ST-TrajVis to help people interacting with the combination of 2D map and space-time cube for dealing with some trajectory data. Beside to these main visual representation component, the system contains another data querying feature and a data enhancement area. The role of the 2D map is to help users investigating spatial information and assist them to answer some basic analytical questions. Meanwhile, if the user wants to get in sight into temporal information of movements or have an overall knowledge about the data, they could use the space- time cube to meet their needs. On the other hand, many interactive methods are provided in the system, such as zooming, panning, rotating, rescaling and so on.

TripVista is an application of visual analysis for analyzing complicated trajectory data at a road intersection (Guo et al., 2011). Users could explore the trajectory from three aspects: spatial, temporal and multi-dimensional information through interacting with the interface of the application. This system mainly focus on trajectory from the micro perspective, which is a difference to some other systems. Research on the moving object, its moving activities and trajectory at the road intersection is the key object in the system. As a result, people can effectively extract traffic flow patterns at the road intersection and analyze them.

Huang et al. (2016) introduced another visual analytics application called TrajGraph to analyze spatio-temporal patterns of massive taxi trajectory data by using a new graph model to reveal the urban network structures. Users can discover the streets usage and learn patterns of traffic condition by the help of interactive methods. Andrienko and Andrienko (2009) introduced an approach to extracting meaningful clusters information from large databases by applying an interactive visual interface. Guo (2009) proposed an interactive visualization environment for a large-sized dataset of migration flows in the USA which provided understanding of spatial interaction and patterns. He also recommended that the user interactive methods in the environment still need some optimization.

2.7.4.   Summary

These sections firstly introduced the concept of visual analytics and how and why it could improve the effectiveness and efficiency of analyzing movements data. Thereafter, the important component of visual analytics - interactive methods was discussed. The task-oriented way of selecting proper interactive methods is recommended. Finally, some visual analytics systems or applications that helps different users to explore and analyze spatio-temporal trajectory data are listed and described. Their works reflect that the visual analytics are helpful to explore trajectory data. However, less such visual analytics were applied to show cycling trajectory data and some of them didn't go through the usability testing with actual users.

2.8.   User research and usability

As this research aims at improving visualizations of cycling trajectory data through conducting user research with usability testing with groups of actual users, some related works of user research and usability testing should be known. Andrienko et al. (2002) tested the usability of interactive maps under the environment of CommonGIS. The interactive tools are emphasized in this work to be investigate how they assist people in exploring the data. They tested tools with carefully elaborated tasks by three standards: (1) Learnability --- to test how users can understand the purpose of the interactive tools and how can they learn to use them; (2) Memorability --- to test the users ability to get the acquired skills and utilizing the tools; (3) Satisfaction --- to test the users' attitudes towards the interactive tools. After three rounds of testing, they analyzed the results and made conclusions that users' participation plays a very important role in the testing and developers should introduce the new concepts and geovisualization techniques to users in a suiTable way.

A use and usability evaluation was made by Koua, MacEachren, and Kraak (2006) to evaluate how a visualization environment that combines maps and parallel coordinated plots meets the users' needs of exploring big amounts of spatial data. In the testing, exploratory tasks and a taxonomy of visualization operations play essential roles. As defined in the paper, the major purpose of analyzing geospatial data is to find patterns and relationships of elements that can be used for answering analytical questions about geographic phenomena (Koua et al., 2006). The taxonomy introduced contains the tasks Identify, Locate, Distinguish, Categorize, Cluster, Distribution, Rank, Compare, Associate and Correlate. Users use these activities to explore the data or to complete analytical tasks. The procedure and results are considered as reference to guide testing work in this research . Robinson, Chen, Lengerich, Meyer, and MacEachren (2005) designed interactive tools for visualizing Epidemiology with the combination of usability evaluation. Before the design, they analyzed users' works by interviewing them and studying their works to better understand the domain. Also they keep communicating with users when they are designing ESTAT (application used). In the testing session, they used recorders to get audio and video recordings of the whole session. Users were also thinking aloud when they were doing tasks. Through these works, they got a better understanding of how ESTAT can help users to explore the data. Besides these works, Kveladze (2015) focuses on user research and usability testing of a space-time cube for visualizing movement data. A well planned usability experiment was conducted in a GeoVisual analytics environment. In this research, usability testing of existing trajectory visualization application will be used with the requirements analysis with actual users together to find limitations of the application and users needs of visualizing the trajectory. So the improved solutions could be provided on the basis of results.

2.8.1.   Some user researches made in related to trajectory visualization works

For the cases of trajectory data, some evaluation works were made when producers implemented the prototype. After Huang et al. (2016) implemented TrajGraph, a usability testing of its interface was also made by conducting user studies. They mainly tested the visualization results by inviting participants to compare information of different regions and to answer questions. For example, they circled two regions and asked participants to find which region is more often used by taxi drivers in a particular time interval. And then, they met an urban geographer with a PhD degree to evaluate the application from a professional level. He evaluated and pointed out that the application could help users to visualize taxi trajectories but it needs some improvements and adjustments to meet actual users needs. So then they concluded these testings reflect that TrajGraph could effectively support users to explore the trajectory data and have a deep understanding of complex traffic situations in a large city.

Gonçalves et al. (2013) executed user studies with 5 users to test its usability of the system and collected their feedback and attitudes from participants. They formulated some user tasks to test the system for different purposes: testing the temporal search feature, testing different visualizations, testing the identification and comparison ability and so on. From the testing, participants showed a higher preference for the Space-Time Cube than for the 2D map and they indicated that they used the STC more to complete tasks. In addition, they commented that they need more filtering and interaction techniques in the visualization.

Similarly, after the implementation, Guo et al. (2011) also tested their system to collect some user feedback. Domain experts who had knowledge of intelligent transportation were invited and gave a good evaluation of interactive functions of the prototype which largely supports people to explore the traffic trajectory data and detect the spatial patterns in the intersection.

2.9.   Conclusion

Though reviewing literature, some concepts and related works of this research are known. The trajectory can be represented by different ways. Four current visualizations methods of the movement data have their advantages and disadvantages. A selection of them should be considered and made before applying them to visualize data. The visual analytics was developed to help people analyzing spatio-temporal data in which interactive methods are every useful to manipulate maps and complete analytical tasks. User research is essential before, during and after designing the visualizations which have been applied to many works. The next chapter will discuss users research and usability methods and techniques used in the my research in order to help me conducting the research and answering research questions.

3. METHODOLOGY

3.1. Introduction

This chapter describes the methods that will be applied to execute this research. It starts with an introduction of user-centered design and usability testing which is most important in the use research part of the research. Then the methods and techniques which will be selected and used to conduct the user research are introduced. And the concrete work-flow of this research is describe in the next section.

3.2. User-centered design and usability evaluation

In this research, a user-centered design (UCD) approach is used. UCD is a procedure combined with several interactive activities with users which helps designers to investigate, acquire and understand these user's requirements and interests deeply. Donald (1988) defined user-centered design as 'a philosophy based on the needs and interests of the user, with an emphasis on making products usable and understandable'. He also describes that products are usable and understandable when users can know what to do and can tell us what is going on. The core of UCD is the involvement of actual users. Their activities influence processes, directions and results of every step of the whole design. In the initial steps of the design process, users should be involved with prototypes and their interactive activities with the prototypes should be observed and recorded (Corry, Frick, & Hansen, 1997). Then designers should improve the problems and limitations that are detected by users and invite them to evaluate repeatedly.
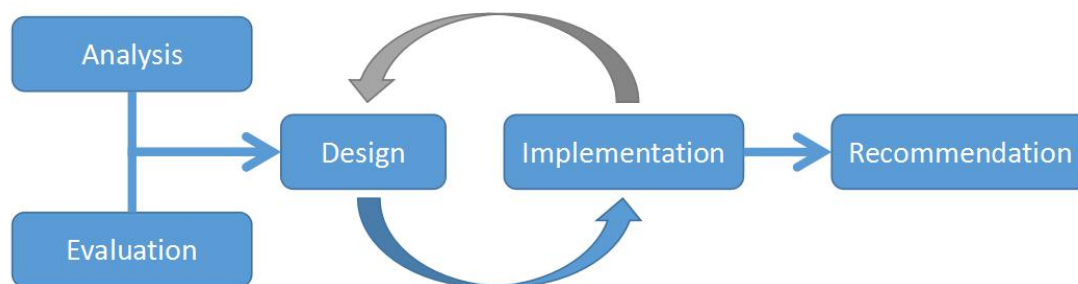
Figure 3-1 The flow of User centered design of this research

Normally, UCD has four main components: the requirements analysis, the conceptual design, the prototype implementation and the evaluation. Users requirements should be analyzed at first. Because this research focuses on the existing visualization application --- BikePRINT which is not useful as expected, the usability evaluation will be conducted to the visualization to find its drawbacks with respect to assisting users answering analytical questions in the first stage, which is an important part of requirement analysis. Usability evaluation is an essential approach in the UCD process. it could be used in all steps of the different products' life cycles. Therefore, I propose an adjusted work flow that is more fit in to this research (Figure 3-1). After the requirement analysis, the conceptual design of improved visualization solutions will be made according to results of previous two parts. Because the developers of BikePRINT are developing the new version of the application, the implementation of designs are some illustrations with explanations as recommendations for them.

For the usability testing, there are five components of usability testing defined by Dumas and Redish (1999):
1) The main goal is to improve the usability of products.
2) Actual users participate in the testing.
3) Users do real tasks.
4) The tester observes and records what users think, say and act.
5) The tester analyzes the recorded result and give recommendation or make improvements

The result of the usability testing should be described by formal reports that contain the evaluation objectives or context, methods applied, the process of testing, result analysis and the summary of the result. The usability could be measured from aspects of effectiveness, efficiency and satisfaction. According to Frøkjær, Hertzum, and Hornbæk (2000) and Bevan, Carter, and Harker (2015), the effectiveness is the standard for measuring the accuracy and completeness of achieving the goal, efficiency is a measurement of resources (time costs, people effort and etc.) that are consumed to achieving a specific goal and the satisfaction is defined as "freedom from discomfort, and positive attitudes towards the use of the product" by ISO 9241-11 which can be used to evaluate easiness and opinions of users while they use the product.

3.3. Methods used in the user research

As the usability testing of BikePRINT will be conducted in the research, a combination of several different evaluation methods are applied to evaluate it. In Kveladze (2015)'s PhD thesis, some methods and techniques and their objectives and advantages are briefly introduced (see Figure 3-2 and 3-3) which could be used in this research according to two aspects: quantitative and qualitative. Using qualitative methods aims to capture and analyze the participants' activities, behaviors, experience, options, etc. and quantitative methods are used to generate mathematical and statistical results of testing data. In this research, questionnaire, thinking aloud, focus group, video/audio/screen recording are used to help me evaluating how do users effectively, efficiently and satisfyingly use the product and exploring what are their needs.

According to Kveladze (2015), using questionnaire is a fast and easy way to know information of a large number of users. Questionnaire is to invite users to answer a list of questions related to the evaluation of products. Before distributing the questionnaire, the tester should prepare it carefully and clearly that will avoid participants' confusion about it. Questionnaire could be used before and after the usability testing. Before the testing, users' profiles, backgrounds, experiences and relationships with the tested product and others could be collected by the questionnaire. After the usability testing, the questionnaire could be used to acquire participates' opinions, experiences, attitudes etc. of the testing procedure. In this research, I will also use Questionnaires to collect users requirements. The questionnaire includes some parts to collect: (1) participants' background, (2) their experience in urban or cycling analysis, (3) their experience and purposes in using 'Fietstelweek' data, (4) their experience in using the web application BikePRINT and results, (5) their evaluation of provided analytical questions and their own questions with respect to 'Fietstelweek' data, (6) their evaluation of provided visualization alternatives.

Thinking aloud is commonly used in usability testing. People use it mainly for collecting participants' thoughts by asking them to speak aloud whatever they are thinking in their mind that is related to their interactions with the task and the product (van Elzakker, 2004). Using thinking aloud can give the tester a deep understanding of cognitive processes. However, it is a little bit time consuming when collecting and

analyzing the data. In this research, Thinking aloud will be used with the combination of screen recording. Screen recording records the user's activities on the screen that reflects how users interact with the visualization system to explore the data. With the help of this method, in this research, users' activities of doing tasks under the environment of BikePRINT can be collected. The focus group is a useful way to collect users experience, opinions and attitudes. People are welcome to share their thoughts to each others. The advantage of it is the focus group is relatively easier and quicker to organized. The procedure of focus group will be recorded by video/audio recorders.

Method	Objective	Advantage	Disadvantage
Focus group	To identify a range of opinions on specific issues or seek community norms	• Group interaction provides a range of issues, opinions and information quickly • Discussion provides detail, justification and clarification • High emancipatory effect	• Less depth of information and less suitable for personal experiences • Managing group dynamics • Hard to analyse
Post-experience Interview	To identify an individual's perception, feelings, opinion and experiences	• Gain in-depth information • Identify context and personal experiences • Useful for sensitive issues • Possible to conduct in many situations	• No interaction or feedback from others • Individual perceptions only • Multiple interviews needed to identify a range of issues
Think aloud	To understand the mental processes in which users are engaged	• In-depth understanding of cognitive processes • Provides rich qualitative data • User thoughts are immediately voiced and memorized	• Cognitive load of problem solving and speaking may be difficult for some users • Time consuming and difficult to verbalize
Observation	To observe how users act and interact in certain situations	• Much contextual and supportive information • Identifies user's actual behaviour • Possible to conduct in many situations	• Interpretation of observations may be subjective • Time consuming and difficult to verbalize
Video recording	To memorize users' facial impressions	• Captures user's emotional behaviour during evaluation to obtain more information • Possible to conduct in many situations	• Interpretation of facial impressions without sound may be subjective • Time consuming and difficult to verbalize
Audio recording	To memorize sessions for detailed analysis	• Allows to extract more contextual information from captured sessions • Possible to conduct in many situations	• Interpretation of user's voice without facial impression may be subjective • Time consuming and difficult to verbalize
Screen recording	To memorize user's manipulation on the screen	• Allows to study and analyse in detail what users did on the screen when interacting with a product • Possible to conduct in many situations	• To understand the reason of particular behaviour on the screen will require user's facial and voice impression • Time consuming and hard to analyse

Figure 3-2 Qualitative research methods made by Kveladze (2015)

Method	Objective	Advantage	Disadvantage
Task analysis	To understand how users perform given tasks to achieve their goals	• Provides rich quantitative data while gathering and organizing information • Focuses on crucial aspects • Obtains knowledge of the tasks that the user wishes to perform • Helps to detect and analyse errors	• Very detailed investigation of complex problems might be time consuming • Requires to develop an approach to measure the results of analysis
Questionnaire (pre-selection)	To generate user's background profile	• Obtains user's professional knowledge and experience regarding the object of evaluation • Easy to process	• Less depth of information
Eye tracking	To understand how and where users interact with the visual environment	• Provides rich, accurate and quick quantitative data • No training is required	• Eye movements of users with contact lenses, long eye lashes, or narrow eyes have gaze recording problems • Users are limited to move or change their position when eye calibration has been done; abrupt movement will result in data loss • Eye calibration requires time

Figure 3-3 Quantitative research methods made by Kveladze (2015)

Before conducting the user research, an elaborate plan should be made beforehand. In this research, the user research starts with preparation and set-up, followed by execution stage and finally to the part of results analysis. In the stage of the preparation and set-up, all needed documents should be prepared. A questionnaire will be sent to participants in advance to collect their backgrounds, experience and wishes. In the stage of execution, some analytical tasks will be used for users to do. Tasks will be formulated according to results of the questionnaire that are closely related to the analytical questions users are interested in. By this mean, when they interact with BikePRINT to complete these tasks, some limitations of BikePRINT that cannot support them in answering their analytical questions could be found. The whole procedure of the testing will be recorded by Camtasia (URL 4). Camtasia is a software which can record testing persons' video, audio and their activities on the screen simultaneously. Then, the relative and important recordings will be transcribed to texts and are then analyzed. When playing the audio recordings, unnecessary contents and unimportant details will be ignored to concentrate on relevant parts associated with the research questions. Then, for analyzing transcribed text documents, the qualitative research software ATLAS.ti (URL 5) will be used to code different text segments to identify different usability issues of BikePRINT. Codes are used here to represent different meanings which could be used to classify a large number of textual or other data units. Then, improved visualization solutions will be made on the basis of the analysis of testing results.

3.4. Work flow set-up
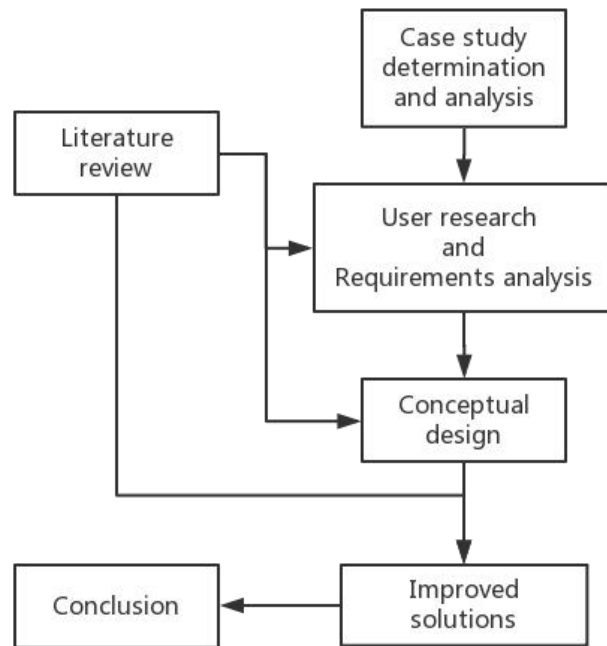
Figure 3-4 Flowchart of research plan and methodology

3.4.1. Literature review

The reviewing of literature of related background and works should be the first step before all other works. In this step, the knowledge of spatio-temporal cycling trajectory data, geovisualization approaches, geovisual analytics, interactive techniques and user and usability studies should be learned. Characteristics of cycling trajectory data and the ways of visualizing the data should be reviewed. The remaining problems and challenges of using visualizations of exploring trajectory data should be described. Moreover, I will get insight into visualization solutions of other sources to find whether there are any potential solutions or ideas can be used in improving visualization of trajectory data. The outcomes of this step are in Chapter 2 of this thesis.

3.4.2. Case study determination and analysis

In this step, the case study is decided firstly. This research focuses on cycling trajectory data as collected in the 'Fietstelweek' and its visualization application BikePRINT. 'Fietstelweek' was initiated to gather the usage data of cycling as much data as possible in the Netherlands. The dataset could be used to analyze spatio-temporal patterns of cycling trips, discover the individual bike user's behaviors or evaluating the traffic safety and road usage. BikePRINT, as one of the visualization solution to the 'Fietstelweek' data, will be evaluated to find whether it has some drawbacks that can be improved or not. Users focused in this research are employees in some municipalities in the Netherlands whose works are related to analysis of cycling because they are actual users of 'Fietstelweek' data and the application BikePRINT. Before the user research, self analysis of BikePRINT will be made to find some limitations out. In Chapter 4, contents about the case study will be presented.

3.4.3. User research and requirement analysis

So the user requirements analysis will be made with the actual user groups. The main purposes of this user research are discovering: 1) what spatio-temporal questions do they have on the basis of 'Fietstelweek'

data, 2) with BikePRINT, what questions they cannot solve properly and 3) what problem did they meet when working with BikePRINT. One main part is testing limitations of BikePRINT. The selection and combination of user and usability methods will provide comprehensive and meaningful results (Kveladze, 2015 #123). In this research, some methods will be used in the procedure including questionnaire, thinking aloud, screen, video and audio recording and focus group. Participants will be invited to work with BikePRINT and will be provided different tasks to do. I will observe and record how do they interact with the application and whether they can solve their problems effectively, efficiently and with satisfaction. Then, Focus Group discussions will be made to collect their experience and opinions about the testing and needs for improved visualization solutions. In this stage, questionnaires will be used to establish the characteristics of the test participants and collect some of their interested analytical questions. The whole procedure of these works will be described in detail including the set-up, preparation, the implementation, the results and analysis in chapter 5.

Parallel to the procedure of the user research, the evaluation of BikePRINT will also be made from my respective. The outcomes of this work could be the hypothesis that will be tested and also be presented to users in the user research implementation. At the same time, potential improved visualization solutions will be collected that could be provided to used to evaluate in the user research implementation.

3.4.4. Conceptual design and Improved solutions

When finishing the user research step and getting results of the user research, conceptual design of the visualization solution will be developed on the basis of results of the user research and my own evaluation of BikePRINT. Andrienko et al. (2003) and other researchers have introduced some guideline for choosing appropriate visualization tools and designing exploratory techniques of spatio-temporal data, which is useful in this research to do the conceptual design. Finally, the recommendations of improvements will be summarized. In chapter 6, these designs will be presented.

3.4.5. Discussion and Conclusion

In the end, the discussion and conclusion will be done in the thesis writing.

3.5. Conclusion

In this chapter, methodologies which will be used in the research are introduced. Making use of these methodologies will assist me tackle problems and answer research questions.

4.   CASE STUDY: 'FIETSTELWEEK' DATA AND THEIR ANALYSIS THROUGH BIKEPRINT

4.1.   Introduction
In this chapter, firstly the basic background of 'Fietstelweek' data and BikePRINT will be introduced. Then characteristics of 'Fietstelweek' data will be described followed by related analytical questions. Next part is the description of BikePRINT and its visualizations. As BikePRINT has some possible limitations, a self evaluation was made to test it, which will be also introduced.

4.2.   Backgrounds of 'Fietstelweek' and BikePRINT
As one of the countries that has a very large group of cyclists, the Dutch government has invested much in optimizing cycling infrastructures. Meanwhile, Dutch municipalities and provinces increasingly value cyclists. Therefore, policy makers and planners in these organizations want to get better insights into bicycle behavior and network performance which could make these investments more effective. They need a cycling usage dataset and knowledge of the current bicycle use to support exploring and analyzing road network quality and cycling behaviors in the current situation, which, however, are lacking. So that they could make policy to optimize cycling environments and cyclists' experience. To reduce this gap between data, knowledge and cycling policy, the 'Fietstelweek' was initiated to collect cycling usage data and BikePRINT (URL 2) was developed to help translate these data into policy-relevant data.

4.3.   Characteristics of 'Fietstelweek' data
The 'Fietstelweek' is the biggest bicycle survey in the Netherlands that accumulates data and insights in how to facilitate all cyclists in the Netherlands. The purpose of launching 'Fietstelweek' was to gather as much data as possible about the bicycle trips in the Netherlands and to identify the bicycle network. Approximately 38,000 people in the nation have actively participated in the this project and provided 1.2 million registered bike rides in 2015. So, the collected data 'Fietstelweek' cannot represent the actual number of bike rides, but only the rides of those volunteers who participated. The first 'Fietstelweek' took place in the third week of September 2015 and will be repeated annually. Employees of the project collected the raw GPS points data which describes trajectory of cycling trips. Then they also processed and aggregated data into different formats that could be downloaded from the website (URL 6). From the website, we can have a brief understanding of characteristics of 'Fietstelweek' data. There are two types of dataset could be downloaded: (1) raw GPS data and (2) processed data. I introduce it one by one in following sections.

4.3.1.   Raw GPS dataset of 'Fietstelweek'
An official document of data description could also be downloaded along with the dataset. Some description of the characteristics of the data are made with the help of this official document.
When I opened the Table of data file, as Table 4-1, some columns that describe the data mean :
1)   ID: Each row represents one recorded GPS point. Rows that have the same ID belong to the same one cycling trip so they are made by one cyclist at once.The first point is the origin and the last point is the destination of the cycling trip.

2) Time: represents the date and time when this point was recorded at that location.
3) Lon and Lat: the location of the point in WGS84 (EPSG 4326).
4) Speed: It is the instant GPS speed of the cyclist at the location.
5) Heading: It is the direction of the cyclist at the location.

Table 4-1 is a part of the Table of data file. As we can see, in the first column 'id', 89, 138 and 28 are route id. Take 'id' = 89 as the example. There are 12 rows have this id which means they constitute a cycling trip '89'. From row 2 to 13, the second column 'time (tijd)' changes from 2015/9/14 8:44 to 2015/9/14 8:47, which means the trip started from 8:44 on day 14, month September, year 2015 to 8:47 on day 14, month September, year 2015 whose duration (traveling time) is about 3 minutes. When plotting these 12 points in 'Google Earth' (Figure 4-1), people can see a trajectory of this trip but it is not continuous.

Table 4-1 Screen-shot of file 'routes1.csv'

	A	B	C	D	E	F	G
1	id	tijd	lon	lat	speed	heading	
2	89	2015/9/14 8:44	5.077098	51.542231	0	0	
3	89	2015/9/14 8:44	5.077239	51.542309	16.164	355	
4	89	2015/9/14 8:45	5.077333	51.542948	18.431999	7	
5	89	2015/9/14 8:45	5.077444	51.543539	17.675999	6	
6	89	2015/9/14 8:45	5.077552	51.543713	13.716	26	
7	89	2015/9/14 8:46	5.077296	51.543822	12.204	256	
8	89	2015/9/14 8:46	5.076313	51.543882	17.352001	273	
9	89	2015/9/14 8:46	5.075352	51.543944	18.143999	274	
10	89	2015/9/14 8:47	5.074415	51.543932	18.036001	268	
11	89	2015/9/14 8:47	5.073444	51.543932	17.459999	270	
12	89	2015/9/14 8:47	5.072414	51.543915	16.416	270	
13	89	2015/9/14 8:47	5.071434	51.543897	16.632	269	
14	138	2015/9/14 19:21	5.077544	51.544875	20.880001	185	
15	138	2015/9/14 19:22	5.077341	51.544214	14.903999	195	
16	138	2015/9/14 19:22	5.077129	51.543513	23.219999	184	
17	138	2015/9/14 19:22	5.077049	51.542827	23.976	185	
18	138	2015/9/14 19:22	5.076949	51.542082	23.508001	185	
19	138	2015/9/14 19:23	5.07686	51.541419	21.203999	185	
20	138	2015/9/14 19:23	5.076716	51.540762	14.184	137	
21	138	2015/9/14 19:23	5.077376	51.540747	17.459999	94	
22	138	2015/9/14 19:23	5.078311	51.54071	19.512001	94	
23	138	2015/9/14 19:24	5.079085	51.540842	12.816	187	
24	138	2015/9/14 19:24	5.078487	51.540799	8.6040001	219	
25	28	2015/9/14 9:56	5.081478	51.544756	15.408001	294	
26	28	2015/9/14 9:56	5.080762	51.544808	3.5999999	214	
27	28	2015/9/14 9:57	5.08036	51.544091	0	0	
28	28	2015/9/14 9:57	5.080465	51.544242	0	0	
29	28	2015/9/14 9:57	5.081156	51.544353	5.6880002	85	
30	28	2015/9/14 9:57	5.081242	51.54462	4.5	109	
31	53	2015/9/14 20:35	5.076649	51.541229	0	0	
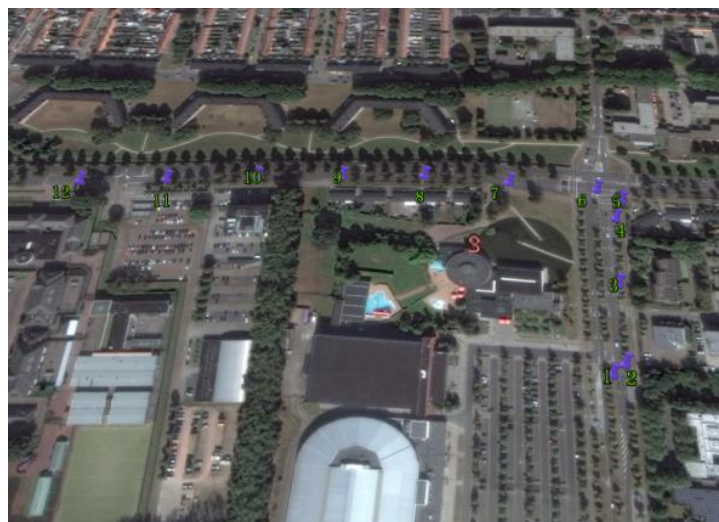
Figure 4-1 GPS points of the cycling trip 'id = 89' shown in Google earth

As introduced in the official document, the sample files have been edited to make it impossible to identify persons. The 'real' raw data do not include these operations:

- Some routes are omitted their real starting and ending points for security issues.
- Some times are moved forward or back accidentally seconds (maximum 60 minutes).
- The codes that identify real Individuals are omitted.

This leads to some inaccuracy about the real conditions of cycling usage. For example, because of the second item, times were changed so that it cannot reflect real patterns of a given time interval. The raw GPS data were also processed into to three types of aggregated datasets: the 'Link' dataset, the 'Button' dataset and the 'Route' dataset, which are the real datasets that BikePRINT visualized.

4.3.2. Aggregated datasets of 'Fietstelweek'

In the section 2.3.2, we introduced that GPS points of trajectories could be aggregated into the existing road network or be used to construct the network. So the trajectory could be represented by a sequence of road segments (links). In the 'Link' dataset of 'Fietstelweek' dataset, links represent road segments where cyclists ride their bikes on. The road segments were made on the basis of OpenStreetMap network and only road segments that were used by cyclists are shown. The 'speed' and 'intensity' information in this dataset are useful for researchers to analyze how people ride bikes and which roads were chosen mostly by them. In the whole country, there are 544288 links that constitute the road network. Loading the dataset in ArcMap generated the display shown as Figure 4-2 and 4-3 and there are some attributes in this dataset (Table 2). The average speed was calculated and could be used to analyze the accessibility of the road network with the combination of the 'point' dataset which provides delay information in the nodes. Calculating transport accessibility could help people to know which is the theoretical fastest way between two locations. Moreover, people also could use it to find on which road segments they can get higher cycling speeds. And the intensity information in this dataset can be used to show the load of each road and cyclists' preferences of choosing roads. Description of columns of Table 4-2 are:

1) FID: the unique ID of each link.
2) Shape: means links are 'polyline' format.
3) Link number (LINKNUMMER): every link has its own unique number. This column is the bridge between 'link' and 'route' datasets. In the datasets, trajectories are represented by this column.
4) Speed (SNELHEID): represents the average speed in the segment.
5) Intensity (INTENSITEI): the amount of cycling trips that were made by cyclists in the road segment in the whole time period of data collection (7 days).
6) Relative speed (SNELHEID_R): the fraction of the relative speed in each segment compared to all segments.

Table 4-2 Part of the Table of the 'Link' dataset

FID	Shape *	LINKNUMMER	SNELHEID	INTENSITEI	SNELHEID_R
0	Polyline	21	18.4673	3	.913053
1	Polyline	110	12.1094	1	.597462
2	Polyline	313	28.8721	4	1
3	Polyline	315	25.3745	4	.979824
4	Polyline	478	16.3505	5	.728894
5	Polyline	539	16.5785	4	.700443
6	Polyline	562	13.8646	7	.648209
7	Polyline	639	18.0881	8	.721765
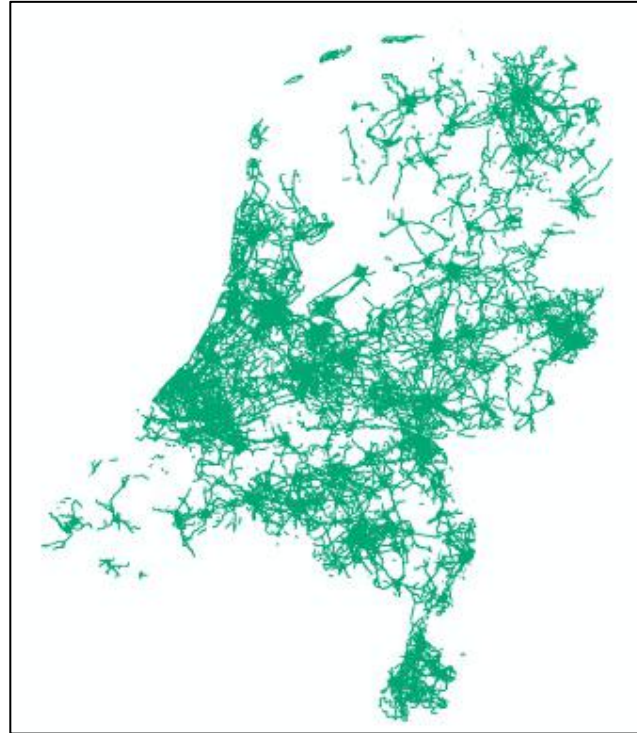
Figure 4-2 Country view of road network in ArcGIS

Figure 4-3 After zoom-in,view of Enschede in ArcGIS. As we can see, the cycling road network are shown by green lines

Another dataset is the 'Point' data containing information of delay points where cyclists stopped and their average delay time. There are 165934 nodes in the whole country. If a cyclist stopped in a location for a time period, the system counts the time he stopped at this location, which means a cyclist was delayed here. If another cyclist also stops here for a while, his/her stopping time was recorded as well. Loading the dataset in ArcMap generated the display shown as Figure 4-4 and 4-5 and there are some attributes in this dataset (Table 4-3). The points were also made on the basis of OpenStreetMap network and they are on

the road segments. The average delay time in each node was calculated with the data of the amount of delay and delay time on this node. There are some columns of information that describe the data:

1) FID: unique number of each node
2) Shape: means nodes are point data
3) Node number (KNOOPNUMME): In the official description document , this column means the location of the node in the network. In this dataset, the location of each nodes are represented by numbers.
4) Time (TIJD): the average delay time per node in seconds of all cyclists who are registered at that node.
5) Number (AANTAL): the amount of cyclists who passed this node in the whole time period (7 days)

Table 4-3 Part of the data Table

FID	Shape	KNOOPNUMME	TIJD	AANTAL
0	Point	-1653292773	1.507962	12
1	Point	43961513	1.621813	7
2	Point	43961968	4.2615	6
3	Point	-312477607	1.961933	7
4	Point	-376377521	4.492036	13
5	Point	1404608321	.740250	6
6	Point	761562561	7.841784	18
7	Point	761562827	2.217385	12
8	Point	30782592	12.889	9

knopen-28992

Figure 4-4 Country view of 'point' dataset in ArcGIS

Figure 4-5 After zoom-in,the delay points in the area of Enschede shown in ArcGIS

In addition, there is another data type 'route' that represent cycling trajectories. In this dataset, a sequence of road segments are used to represent the trajectory of cycling trips. By this dataset, we cannot know which are the exact origin point and destination point of one cycling trip. We only can know which links are the first one or the last one that the cyclist passed through. In this dataset, one row is one item of data which consists of 'routeid', link number, direction, hour and weekday. When I opened the dataset in Excel, the part of Table is show as Table 4-4.

Table 4-4 Part of the Table of 'Route' data

	A	B	C	D	E	F
1	routeid	linknummer	richting	snelheid	uur	weekdag
2	543604536	1550580	t	19. 736647	8	4
3	543608349	1346272	t	17. 393603	7	4
4	543610424	839906	f	15. 083163	8	4
5	543646363	1890915	t	9. 1413642	8	4
6	543646363	1890914	t	7. 481261	8	4
7	543646363	1890913	t	4. 6301499	8	4
8	543657262	392290	f	14. 640018	9	4
9	543657262	391232	t	15. 469775	9	4
10	543657262	1692784	t	16. 917339	9	4
11	543657262	264202	f	13. 199229	9	4
12	543658527	1492563	t	18. 123415	11	4
13	543663229	1756802	t	17. 638158	8	4
14	543686769	1999149	t	15. 740869	19	3
15	543686772	1693365	t	16. 045286	9	4
16	543686772	1108051	t	18. 700658	9	4
17	543686772	1107834	t	20. 517725	9	4
18	543686772	1360016	t	17. 504387	9	4
19	543692197	99413	f	17. 84928	11	4
20	543699536	1443337	f	12. 772295	8	4
21	543703209	106269	t	19. 050396	9	4
22	543703209	115811	t	13. 446265	9	4
23	543720714	767700	f	18. 011382	8	4
24	543721586	1787743	f	22. 004285	7	4
25	543721586	103791	f	21. 138486	7	4
26	543721586	103792	f	21. 510809	7	4
27	543721586	103793	f	21. 100371	7	4
28	543721586	103799	f	17. 033091	7	4
29	543721586	103795	f	17. 039204	7	4
30	543721586	208307	t	19. 875243	7	4
31	543736646	79923	t	16. 616807	9	4
32	543736646	78726	f	15. 823339	9	4

routes

The meaning of attributes in the dataset are:

1) routeid: This column means a unique routeID of the cycling route. If some rows have same one 'routeid', they belong to one cycling route (cycling trip) with this routeID.

2) Link number (linknummer): this column contains same numbers with column 'Link number' of dataset 'Link data', which means these links form part of one route if they have the same 'routeid'.

3) Direction (richting): this column indicates in which direction on the appropriate link is cycled through this route. 'T' and 'F' represent diverse directions of the movement on the segment.

4) Speed (snelheid): represents the average speed in the road segment.

5) Hour and Weekday (uur and weekdag): these two columns mean the time that the cycling trip was finished in each road segment.

4.3.3. Summary

This section describes some characteristics of the 'Fietstelweek' data. In addition to the raw GPS data, it also provides aggregated datasets, which I focused on in this research. Road segments are the base of the dataset when the analysis of trajectory data is made, because a sequence of road segments is used to represent a cycling trip (as shown in Figure 4-6). This way of representing trajectories is also used in BikePRINT. By this mean, it is not easy to analyze the movement from the perspective of inside the trajectory. For example, how the cyclist moves in the trajectory cannot be analyzed. The analysis can be made when dealing with several trajectories. With respect to the data, some analytical questions are proposed to explore and analyze the cyclists' movement.
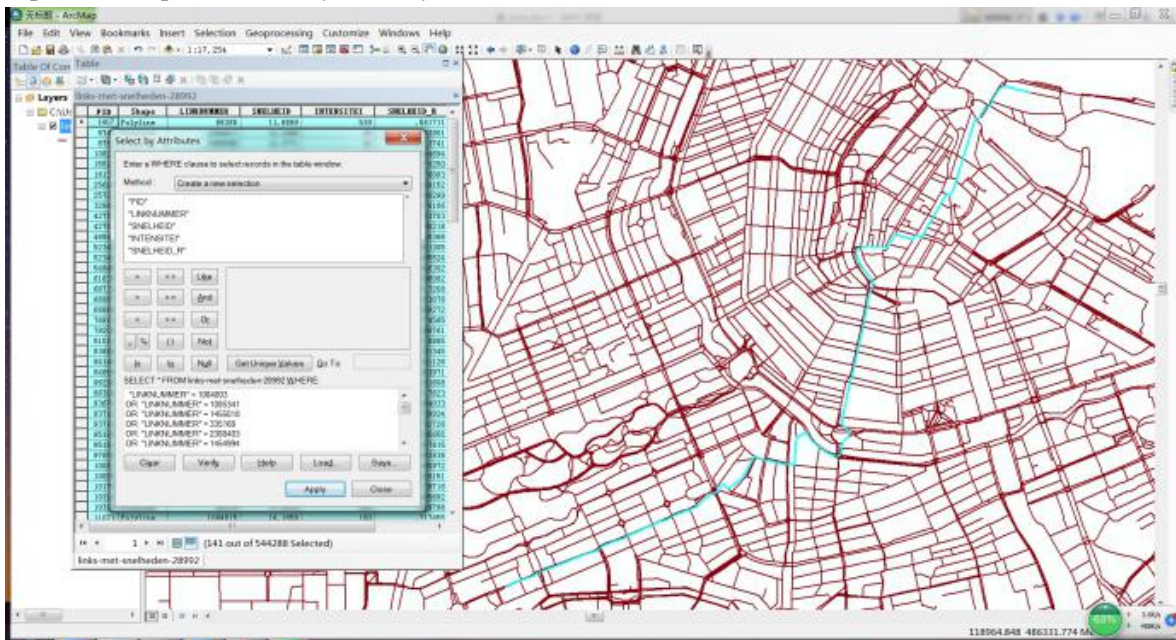
Figure 4-6 Blue lines represent trajectory of one cycling trip in ArcGIS

4.4.   Analytical questions with respect to the data

4.4.1.   Introduction

In this section, some analytical questions are formulated with respect to cycling trajectory data. It is possible for those who are not very familiar with spatio-temporal movement data to have better understanding about data and know what questions could be probably answered. Also, these questions are

may used by users in their job. Moreover, having a good understanding of user analytical questions and tasks helps the designer/producer of data analysis tools to collect user' requirements (Boyandin, 2013).

The way of proposing analytical questions is guided by the taxonomies of questions which will be described below. These taxonomies guide and support people visualizing, exploring and analyzing trajectory data from different levels. For example, if the user wants to explore the spatial information of trajectories when temporal and/or attribute components are given, the question could be: Where are the trajectories made by Peter in day 1?, in which 'Where' is the seeking information, 'made by Peter' is the given attribute and 'in day 1' is the given temporal component. Purposes of making these analytical questions are: (1) providing them to actual users to evaluate and group which are their interests and which are unsolved analytical questions; (2) some analytical questions will be used to evaluate the BikePRINT by myself. Since BikePRINT provides several visualizations rather than only the visualization of trajectories, not only questions about trajectories are proposed here but also questions for answering other types of analytical questions in BikePRINT. For example, questions about road usage were made as well, which are also concerned by urban planners. These questions are grouped to four levels as followings.

4.4.2. Classification of analytical questions and tasks

1) Peuquet triad framework

Peuquet (1994) introduced a triad framework to represent spatio-temporal data that contains three fundamental components: location (Where), Time (When) and Attribute (What). Users' analytical questions can be collected and classified according to these three types:

- **When + Where → What**: Describe the moving objects or series of objects (what) that are in a given location or series of locations (where) at a specific time or series of times (when).

- **When + What → Where**: Describe the location or series of locations (where) occupied by a given object or series of objects (what) at a specific time or series of times (when).

- **Where + What → When**: Describe the time or series of times (when) that a given object or series of objects (what) occupied a particular location or series of locations (where).
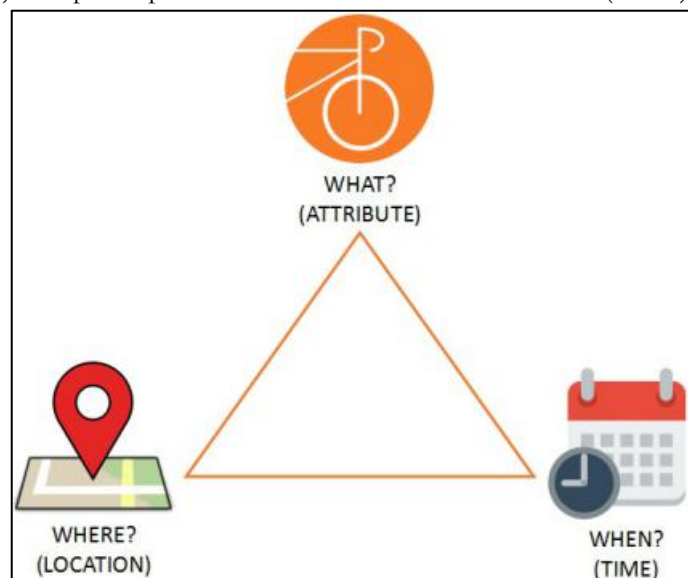
Figure 4-7 The sketch of relations of location, time and attribute

Figure 4-7 is a sketch drawn for showing relationships of the three basic components of the triad framework that gives a good guideline to users to formulate the basic questions so that they can get better insight to the nature of spatio-temporal trajectory data. For example, if the user wants to know the

location or series of locations when the time (when) and attribute (what) components of the object are given, he/she can propose the question like 'Where is the longest cycling trajectory on day 1?'. In this research, I used this framework to propose some analytical questions about the 'Fietstelweek' data. And also, the framework is used to examine users' interested or unsolved analytical questions and revise them to the right format. On the basis of these analytical questions, corresponding user tasks are made for users to do the the usability testing of BikePRINT.

2)   Bertin's reading levels

This is a taxonomy proposed by Bertin (1983) on the basis of 3 different reading levels:

-**Elementary**: Questions concerning one single element.
-**Intermediate**: Questions concerning a group of elements.
-**Overall**: Questions concerning all the elements together.

Some analytical questions could be asked by different levels on the basis of this theory. In cycling trajectory data, questions about the single movement are elementary questions, questions about a group of several movements are intermediate questions and questions about the all element are overall questions. In this research, since the data type we focus on is the spatio-temporal trajectory data, temporal information is also indispensable to consider. Therefore, questions concerning the temporal elements are also important to deal with, called temporal questions. Then, the combination of methods of Peuquet triad framework and Bertin's reading levels is helpful for us to propose some analytical questions with respect to 'Fietstelweek' data. For example, the question 'Where is the longest trajectory?' is an elementary question focusing on the spatial component (Where) of the trajectory.

3)   Tasks taxonomy of Shneiderman

Shneiderman (1996) made the taxonomy with the consideration of designing user interfaces of visual representation which called the 'Visual Information-Seeking Mantra' that consists of seven tasks:

-**Overview**: Get an overall look of all entities.
-**Zoom**: Zoom in to the targeted item(s).
-**Filter**: Filter out unneeded things.
-**Details-on-demand**: Select interested/needed information to have detailed look.
-**Relate**: Investigate relationships among items.
-**History**: Keep a history of actions to support undo, replay, and progressive refinement.
-**Extract**: Allow extraction of sub-collections and of the query parameters.

He also listed seven types of data that these tasks can be applied to: one-, two- and three-dimensional, temporal dimensional, multi-dimensional, network and tree data.This taxonomy is useful to guide people effectively exploring the data from the overall perspective to the details.

4)   Andrienko's theory of taxonomy

Andrienko and Andrienko (2006) also introduced a taxonomy for spatio-temporal data that could be used to deal with the movement data in this research. One idea of this taxonomy came from Bertin (1983)'s taxonomy. The difference is that Andrienko and Andrienko (2006) classified tasks into two main levels:

- **Elementary tasks**: Focus on individual elements of the reference set.
- **Synoptic tasks**: Focus on the reference set as a whole or its subsets as wholes.

And there are some sub-tasks in this two types of tasks:

For Elementary tasks :

1) Look-up tasks: to find values of some components that contains the given value.
2) Comparison tasks: to compare values or characteristics between different elements or between elements in different conditions or given values. For example: Please compare the health conditions of a cyclists when he/she was in starting point and ending point of the trajectory he/she just made.
3) Relation-seeking tasks: to investigate relations of characteristics between elements or group for elements.

For Synoptic tasks:
1) Pattern definition tasks: to ascribe a pattern to a specific type and summarize characteristics.
2) Pattern search tasks: to find patterns similar to the given one.
3) Patterns comparison and relations seeking tasks: to find similar patterns and group them, to compare different patterns or to investigate relations among them.

4.4.3. Proposed questions

- **Elementary questions:**
  1. What is the average number of cyclists on road segment X per day?
  2. What is the average number of cyclists in city X per day?
  3. Where do cyclist ride bikes in city X on day Y?
  4. Which road segments are passed by cyclists per hour?
  5. Which train stations did most cyclists depart from on day Y?

- **Intermediate questions:**
  1. Where is the longest cycling trip in city X on day Y? Where are its origin and destination?
  2. Where is the earliest/latest cycling trip in city X on day Y?
  3. Where is the cycling trip that lasts for the longest time in city X on day Y?
  4. There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?
  5. Where are the cycling trips whose origins are in city X and destinations are outside that city X?
  6. Where are the destinations of cyclists from the central station of city X during the morning rush time of day Y?
  7. How many people rode bikes from city X to other places outside that city on day Y?
  8. Where is the intersection that cyclists need to wait for the longest time?
  9. Which road segments are chosen by cyclists most often on different days in city X?

- **Temporal questions:**
  1. What is the trend of movements from central station of city X to different destinations from morning to night on day Y?
  2. What is the trend of movements between city X to city Y during the whole period?
  3. There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. What is the trend of cycling movements between the central station and this place?
  4. What is the change of movements on 7 different days in city X?
  5. On which day did most cyclists depart from the central station in city X?
  6. Have the spatial distribution patterns in morning rush times of 5 working days changed in city X?

- **Overall questions:**

1.What are the spatial-temporal patterns of cycling trajectories from the east to west part in city X on day Y?

2.What is the spatial-temporal distribution of cycling trajectories in the city which have the highest average speeds on day Y?

3.What are the spatial-temporal patterns of trajectories from city X to city Y in the morning rush time and night rush time of working days? What is the difference?

4.What is the similarity of spatial-temporal patterns of bike trajectories on the 5 working days in city X?

5.What is the difference of spatial-temporal patterns of movements in city X between working days and weekends?

6.What is the difference of spatial-temporal patterns of movements in city X between the morning rush time and night rush time on day Y ?

7.What are the spatial-temporal patterns of cycling trajectories in the city X during the whole period?

4.4.4. Summary

These questions are relatively general. More specific questions provided to a group of users in the user research will be made according to the place where these people focus on. Questions are also used to formulate some tasks for users to do for testing BikePRINT in the user research stage. The next section will introduce BikePRINT and then some questions will be used to evaluate whether they can be answered by BikePRINT or not and how from my perspective.

4.5.   BikePRINT and its visualizations

4.5.1.   Introduction

The next section will firstly introduce some general information of BikePRINT. Then the following one will describe visualizations provided by BikePRINT and then use some analytical questions from the list mentioned above to test whether they can be answered by BikePRINT or not. The possible factors that lead to the limitations will be proposed depending on the testing results. Finally, I will summarize some of its limitations from my perspective and give some alternatives solutions about it.

4.5.2.   General information of BikePRINT

On the basis of the 'Fietstelweek' data source, BikePRINT, a web visual analytics application, was developed by an team of NHTV lecturers and researchers who specialize in issues of Mobility and Spatial Planning which visualizes these data by several web pages for analytical purposes of different policy makers and planners. It aims to meet the need for knowledge of policy makers in the field of cycling and offers supports in policy making and in implementing cycling policy. The web application BikePRINT provides information about network performances of cycling, such as: heat maps of road intensity, delays at junctions, route choice (and detours), average/relative speeds and the number of cyclists per day. It anticipates the need for knowledge of policymakers in the field of bicycles, and serves to provide them with information of people's cycling behaviors and condition of roads usage and gives support to policy planning and executing bicycle policy. BikePRINT translates GPS data of bicycle movements into policy-relevant information and gives users a detailed picture of up-to-date bicycle use, network quality and potential bicycle accessibility. The base map of the application is OpenStreetMap (URL 7) and CartoDB (URL 8) was used to visualize the data. OpenStreetMap is one example of volunteered geographic information. It supports registered volunteers to contribute different types of geospatial information on this platform. CartoDB is an online cartography service system that provides GIS and web mapping tools

for users to visualize geo-based data on maps through an intuitive display. The next sections will firstly introduce visualizations provided by BikePRINT and then I will use some analytical questions to test whether it can help me find answers well or not.

4.5.3.   Visualizations of BikePRINT

In the website of ''BikePRINT, there are two main pages that visualize the data: (1) Bicycle use, and (2) Speed. In each page, there are some layers with the spatio-temporal information. Without log-in account, the data are only provided for a little part of the Netherlands but with full visualization information and functionality (see Figure 4-8).
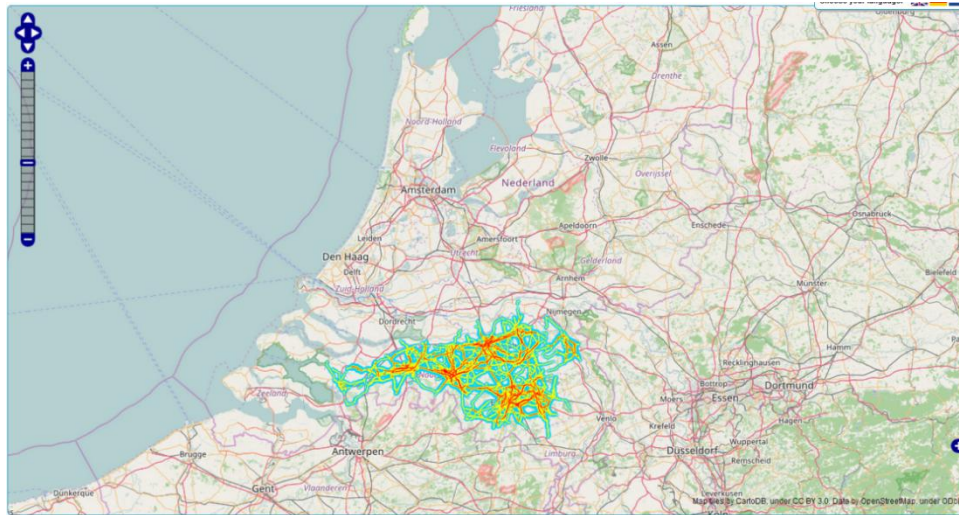
Figure 4-8 Without the full license, the data only can be shown in a little part of area (Brabant) in the Netherlands

We looked at the first visualization page 'cycling use'. In this page, we can move, zoom in and zoom out the base map. In the left part of this page, the data that can be shown are listed in the legend. Users can click layers in the legend to show the corresponding dataset on the map or turn it off. As shown in Figure 13, intensity, routes, intensity shortest route, difference measured/shortest, detour factor, residents, and jobs can be clicked to show. However, layers will overlap each other if the user turn on more than two layers so that a layer will be covered by other one.
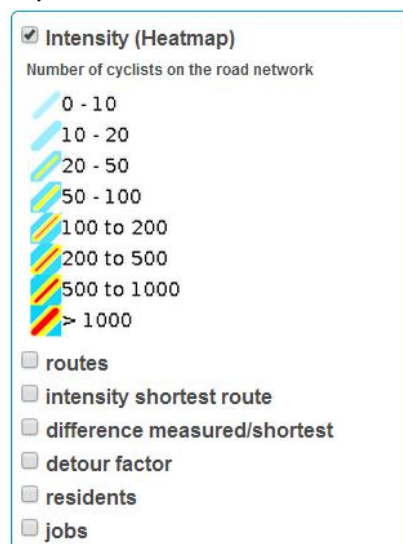
Figure 4-9 The dashboard of turning on/off layers. The legend of visualizations are in this box as well

1) 'Intensity': To get a more clear view, I zoom-in to the urban area of city Eindhoven and chick the button 'Intensity (Heatmap)' in the legend part (left part of this page), the road network with amounts of cyclists on the roads in the whole recording period are shown as Figure 4-10. In the visualization, visual variables of road segments (lines) are used to represent the different values. Colors and sizes (thickness) of lines are used together. Blue thin lines mean that less cyclists ride bike on these road segments and thick lines mean that much cyclists ride bikes on these roads. By adding yellow and red in the line, the makers tried to improve legibility. In this page, planners can know which roads are used commonly so they can improve the quality of roads or make these roads to give cyclists more priorities of cycling.
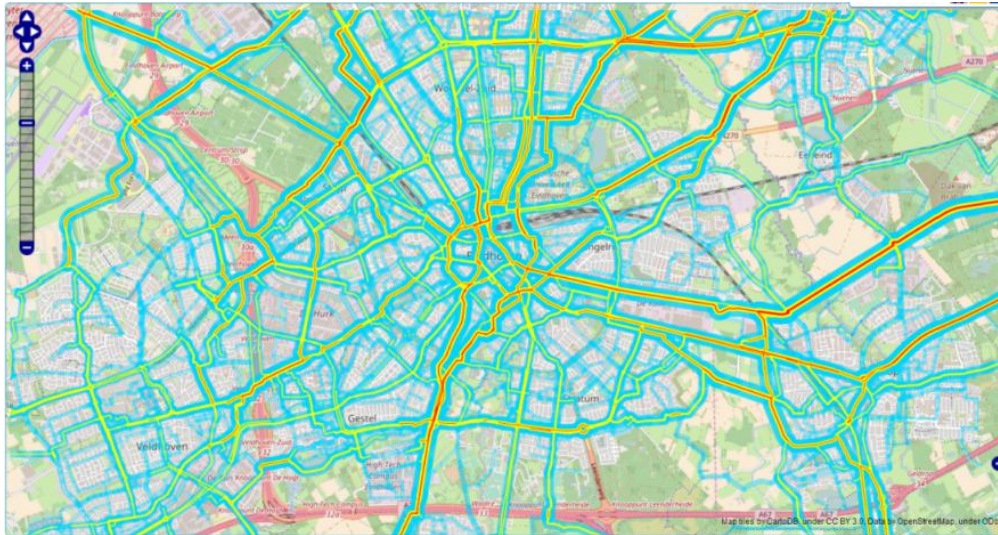
Figure 4-10 The map of main cycling network in the area. In this map, the usage of road network by cyclists are represented by lines with different sizes and different internal colors.

2) 'routes': The second layer in the legend part is 'routes' which shows the trajectory of cycling trip made by cyclists. The thickness of blue lines represent the different amounts of intensity (the number of cycling trips that passed the road segment) which is similar to the previous visualization. However, the emphasis of this visualization is to provide cyclist trajectories. Users can click on a road segment to check which trips passed through this segment. In the map, all unselected routes are shown by blue lines (Figure 4-11).
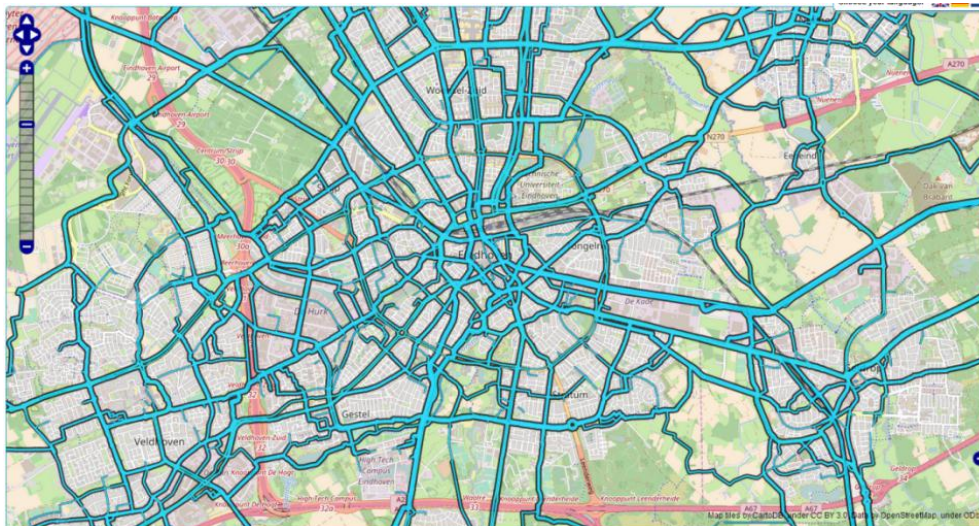
Figure 4-11 The original view of 'route' map. Blue lines represent the road network where cyclists passed by

When the user chooses one road segment and click it, all routes which pass this road segment are highlighted by red (Figure 4-12). However, there are some limitations that bring difficulty to users to explore the information, which will be described in details in the next section.
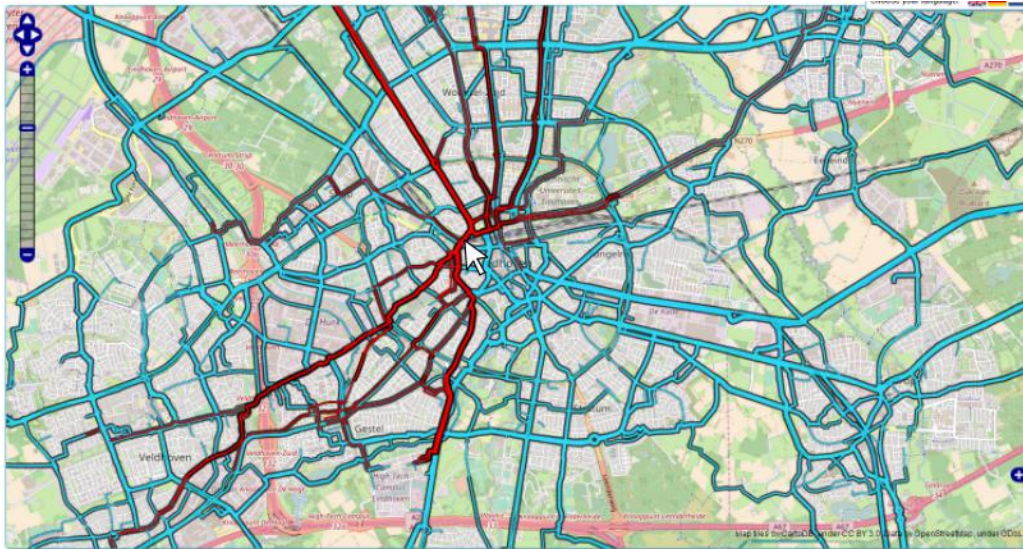
Figure 4-12 Red links are selected routes by clicking a road segment. The mouse points to the location of the clicking but the clicked segment is not highlighted

3) 'Intensity shortest route': fictive number of cyclists if everyone took the shortest route according to a cycle route planner. The different colors and thicknesses represent different amounts (Figure 4-13). And 'difference measured/shortest': the difference between fictive cycle route planner and measured number of cyclists (Figure 4-14). Figure 4-13 shows the difference compared to Figure 4-14. It can reflect how cyclists choose roads to cycle. The blue lines are the positive number of difference which means these roads are more popular. Inversely, the red lines mean that they are less attractive to cyclists. It shows the difference between expected situation and actual situation of choosing the cycling route.
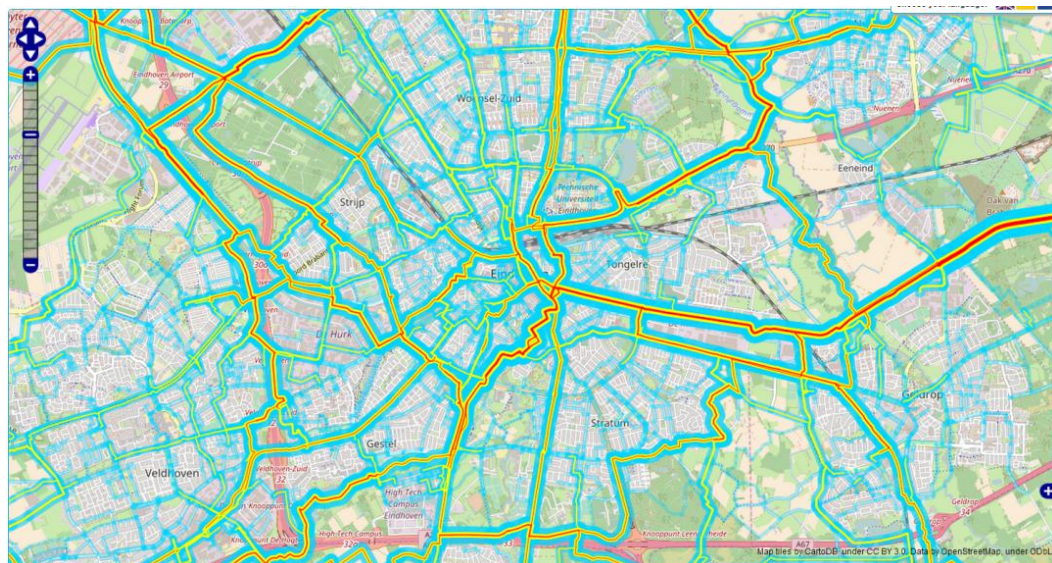
Figure 4-13 The map of fictive shortest routes of cyclists

Figure 4-14 The map of difference between fictive and measured number. Red lines and blue lines are represent inverse values. Thicknesses of lines represent the amount of the value. However, there are visual clutters on the map.

4) Next layer is to show lines with different colors to show how cyclists cycle a detour compared to a straight-line distance (Figure 4-15). The color visual variables of lines are used here to show different detour percentages. The purpose of this layer is similar to the previous one.
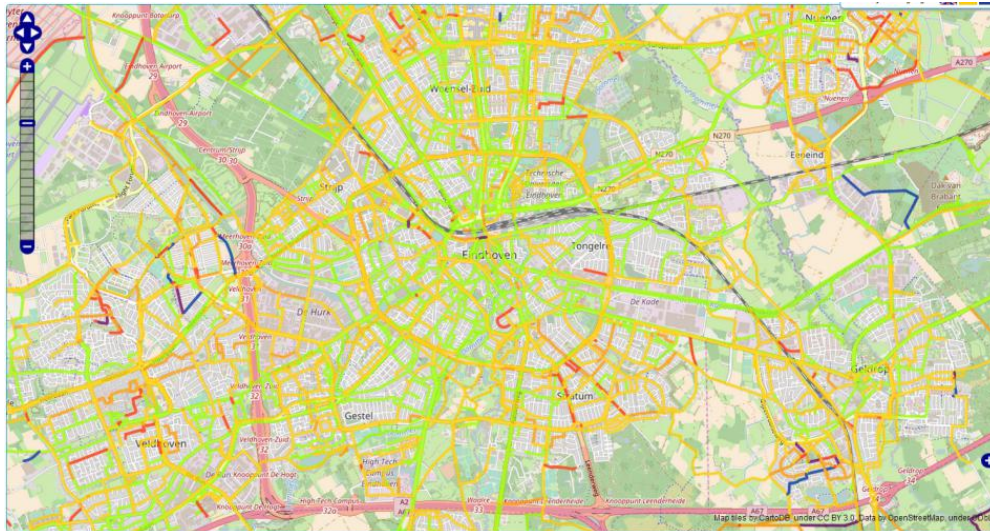
Figure 4-15 The map of detour factor

Another main visualization page, 'speed', mainly focuses on 'speed' attributes including the absolute speed, relative speed and average delay time, which are shown by Figure 4-16 (absolute), 4-17 (relative) and 4-18 (delay). In the legend, colors are used to be visual variables for representing different values.

1) 'Absolute speed' map

In this map, different colors represent different absolute speed interval. The absolute speed means the average speed of cyclists in each road segments.
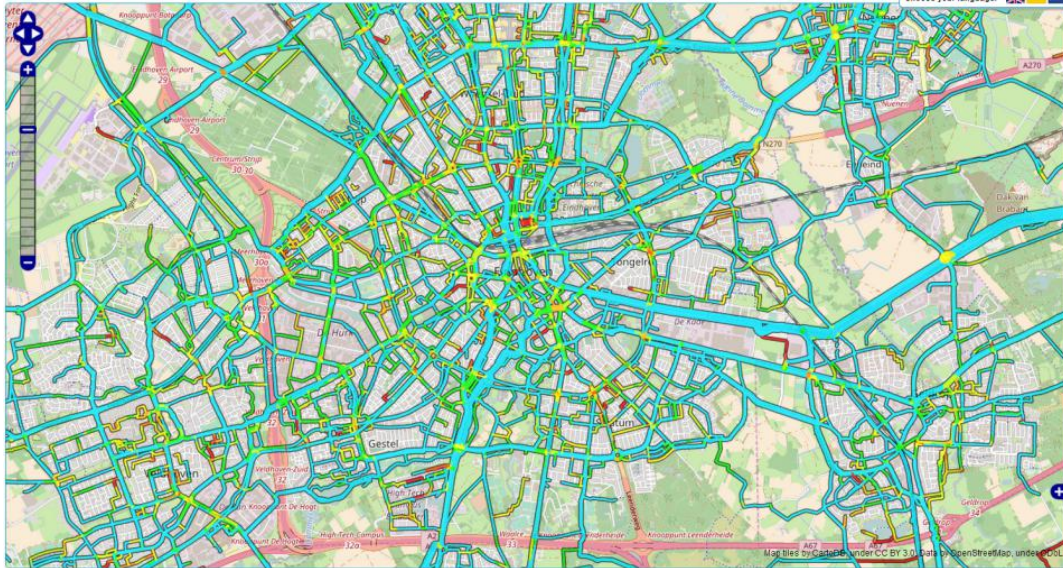
Figure 4-16 The map of absolute speed. Colors of lines are used to represent different average speed

2)   'Relative speed' map

The relative speed is compared to the mean speed of all routes on each road segment. It shows the average speed of a segment compared to the average speed of all segments. Different color visual variables of segments represent different compared values. Traffic lights in the intersection are shown by red which denote bottlenecks. In this visualization, because many values are represented by similar colors, people cannot distinguish them very well.
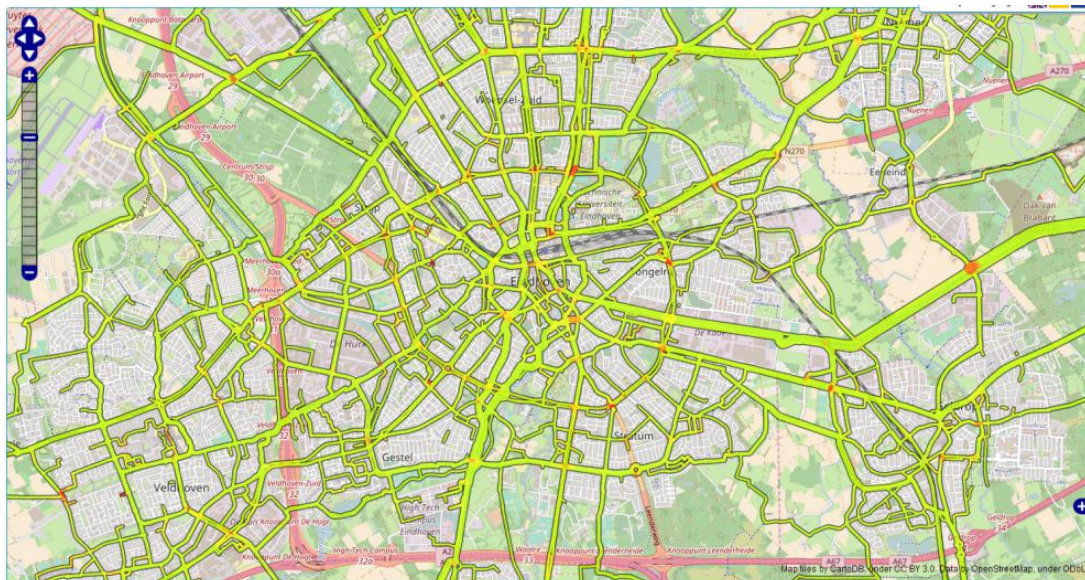
Figure 4-17 The map of absolute speed

3)   'Delay' map

In this map, different colors of points represent different average waiting time intervals in each junction and the size of points means the amount of cyclists who passed the point for the whole period. For some small red points, they mean only a few cyclists have a long waiting time there. From these three views, users can know which roads could provide cyclists with a faster speed, and in which locations cyclists may need to wait for a moment. In addition to these three types of information, users can click on an area or a road segment to see travel times from that area/road (See Figure 22 and 23).
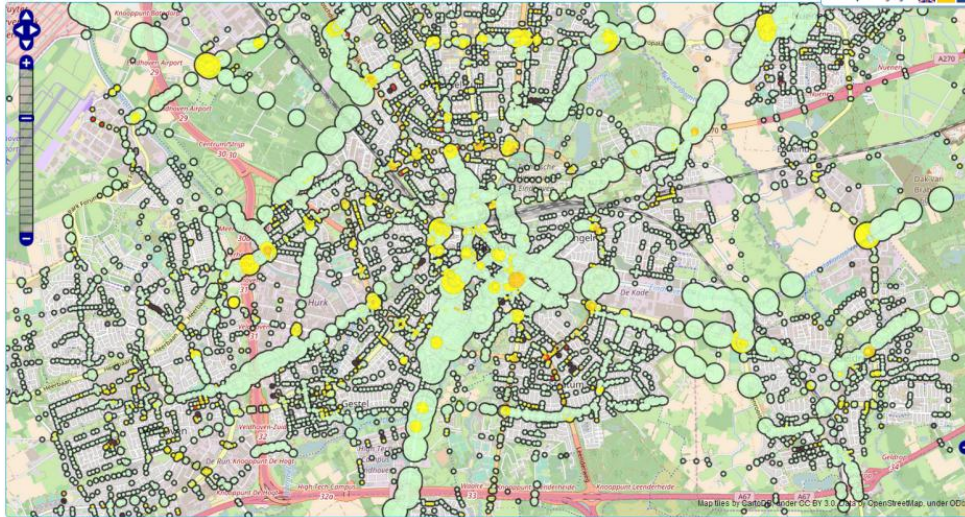
Figure 4-18 Points of delay time. Colors of pints are used to represent delay time and size means the amount of cyclists who had this delay time
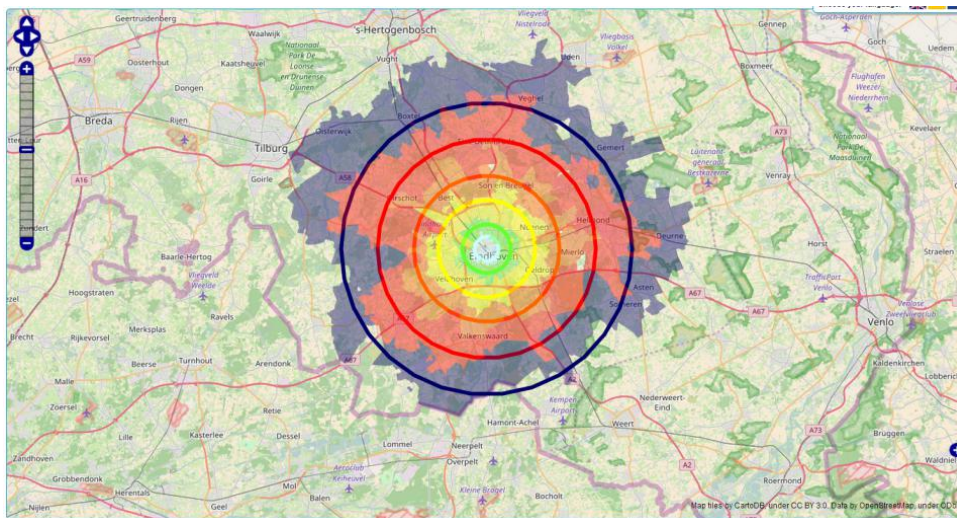
Figure 4-19 Area representation of isochrones. When the user clicks a location, he/she can know how far can he/she go from this location in different times.
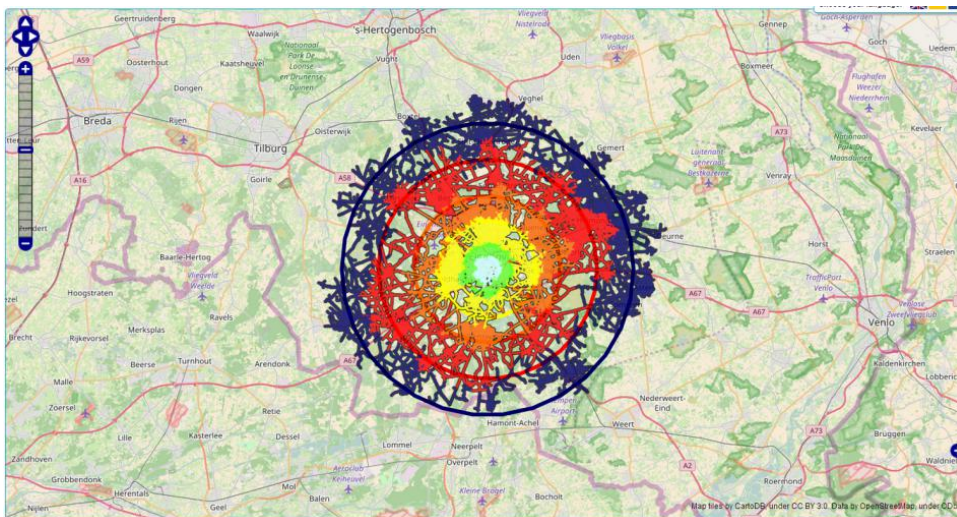
Figure 4-20 Network representation of isochrones. When the user clicks a location, he/she can know how far can he/she go from this location in different times.

4.5.4. Conclusion

From the meeting with some staffs who are involved in the development of BikePRINT, we learned that the responses of users show that the application has some limitations. Through analyzing visualizations of BikePRINT, I found that there are some limitations from my perspective. However, whether it really affects actual users to use it not very well or not is unknown. So, in the next section, a description of self evaluation of BikePRINT is made to test some hypothesis before conducting the user research.

4.6. Self evaluation of BikePRINT

4.6.1. Introduction

User research with usability testing will be made to investigate the limitations of BikePRINT. Before conducting the experiment with actual users, I made a self evaluation of BikePRINT for the following purposes: (1) having a deep understanding of BikePRINT before executing the experiment with actual users; (2) finding some limitations of BikePRINT from my perspective which could be compared to users' findings in the user research; (3) finding out which analytical questions cannot be solved with the help of BikePRINT. The results of this evaluation are described in this section.

4.6.2. Limitations of visualizations in BikePRINT

1)   'Intensity', 'absolute/relative speed', 'detour' and 'shortest route' maps
These maps are used to show attributes of road segments (e.g. Different speeds and amount of cyclists who passed by) so we discuss them together. The amount of each attribute should be represented on the map. In 'intensity' map, the sizes (thicknesses) of blue lines with inner colors represent different amount well that helps users recognizing the different amounts easier. However, in 'absolute/relative speed', 'detour' and 'shortest route' maps, different colors are used to represented amounts which is not very easy for users to read the map if they don't refer to the legend. Moreover, in these maps, some color visual variables of the road network look similar to the colors of the base map, which may cause that users cannot distinguish the information very well (Figure 4-13). Moreover, these maps only show the attributes of road segments for the whole time period (7 days). Therefore, the user cannot extract information of a particular time interval (e.g. to know what happened in the morning). And there is no interactive functionality to filter road segments with particular values (e.g. Intensity more than 1000). Lastly, these information are visualized separately by different layers so that users cannot analyze them together.

2)   'Route' map
In the 'Route' map, no temporal components of movements are represented. The routes that are shown in the map are those for the whole time period (7 days). Trajectories are only highlighted when a road segment that is part of these trajectories is clicked. Therefore, spatial patterns of trajectories cannot be extracted. In the 'Fietstelweek 'dataset, there is information about the direction of movements, but in BikePRINT, the direction is not represented.

So, the only way to select and highlight a trajectory or trajectories is clicking the road segment that is part of these trajectories. There are no other ways for selecting trajectories. Users cannot select trajectories through common methods, like filtering attributes. And when a user clicks a road segment that is part of many trajectories, all these trajectories are shown. Besides, these trajectories always overlap each other so that users cannot distinguish different trajectories (Figure 16). A user cannot filter a single trajectory from them.

3)   'Delay' map
In the delay map, points do not have clear definition of different size and color visual variables. In the legend, only color visual variables are explained as delay time but the meaning of size is unknown. And

many points on the road segments generate some clutter. Besides, these points cannot be selected by users to know detailed information.

4.6.3. Is BikePRINT helpful for answering the proposed analytical questions?

In order to test how these limitations effect people to answer analytical questions, all proposed questions related to the 'Fietstelweek' data in section 4.4 were looked through. Of the elementary level questions, no questions could be answered with the help of BikePRINT very well. For example, to answer questions like 'What is the average number of cyclists in city X per day?', the user has to know the amount of cyclists who cycled in city X on different days and calculate the average number. However, in BikePRINT, information cannot be shown for different days and there is no feature to calculate this number. Of the intermediate questions, only question 4 (There is a place outside of the city center where many students or employees commute to / from every day. Which route is most popular between the central station and this place?) and 8 (Where is the intersection that cyclists need to wait for the longest time?) could be answered with the help of 'delay' map in BikePRINT. For other questions related to trajectories, origin/destination, directions and temporal component are not visualized in BikePRINT so that theycannot be answered. In addition, since there is no temporal information provided by BikePRINT, the questions with a time component cannot be answered. Therefore, the questions in 4.4.3 which are focusing on temporal information (trend, change, etc.) cannot be answered. Finally, BikePRINT cannot provide the features for users to compare different information (e.g. different spatial patterns).

4.6.4. Summary of self-evaluated limitations in BikePRINT

Through looking at BikePRINT in detail, the above limitations of all visualizations could be concluded to these general limitations:

- Data are visualized separately. In BikePRINT, different data are shown by different visualization layers and there are not enough connections between them. For example, users may want to analyze trajectory and speed information together. In BikePRINT, they cannot do this very well. Although layers could be opened together, so that they overlap each other, they generate more clutter than information.

- Some visual representations of information are not desirable and suiTable so that it is not easy for users to distinguish and extract the information (e.g. the color visual variables used in 'speed' map).

- BikePRINT cannot visualize temporal information very well. For example, users cannot derive information for a particular time interval (e.g. morning rush hour). So, BikePRINT is more like a spatial data visualization solution.

- Previous limitations are related to visualizing the data. The fourth limitation is about functionality in BikePRINT which should help users to explore the data. BikePRINT only has basic interactive methods which cannot support users exploring the data well. Only panning, zoom-in/out and clicking methods are provided. Less interactive methods can support them to choose components (e.g. To choose a trajectory or many trajectories, or to choose elements in a particular area) to derive more detailed information. Besides, users' operation of BikePRINT is irreversible and no historical records are made. For example, the user cannot choose two and more elements (e.g. Two road segments or more cycling trips) simultaneously to compare them.

However, the question is whether these limitations really restrict the usefulness of BikePRINT and how do they affect people who want to analyze cycling data. So they should be tested with actual users in the next step. Before conducting the user research with actual users, some possible solutions could be proposed to improve on these limitations. These solutions will also be presented to actual users in user

research step. They may evaluate and choose those solutions which are helpful for them to answer their analytical questions.

4.7. Possible improved solutions

According to the limitations described above, some improved solutions could be proposed from three perspectives:

1) Access to needed data/information

One of main limitations of BikePRINT is that BikePRINT only shows data for the whole time period of recording (7 days) and users cannot extract needed information by spatial, temporal components or attributes. The purpose of this type of solutions is to help users extracting their needed data instead of all data so that they can only focus on and analyze these information. Some functionality of filtering and selecting these information should be be added. Firstly, a time filter should be added in BikePRINT which helps users to extract information for a particular time interval (e.g. from 7-10 am). In the 'intensity' and 'speed' maps, spatial patterns in this time interval could then be extracted and visualized. And in the 'route' map, all trajectories made in this time interval could be shown separately and all together. Secondly, a value filter could be used in BikePRINT for obtaining road segments which have the particular values (speed, intensity, etc.). These ways of accessing data could be used together so that some complicated analytical questions with respect to spatial, temporal and attribute components can be answered.

2) Visualizing accessed data

When the user have accessed their needed data, suiTable ways of visualizing them should be available in BikePRINT, which is the purpose of thief type of solutions. Firstly, it should be possible to visualize different data together so that users can analyze them in conjunction. For example, 'speed' and 'intensity' information related to road segments could be represented colors and widths of road segments respectively in the one visualization layer. Secondly, the base map should be a little bit subdued so that users can focus on the thematic information. A third suggested improvement is about visualizing individual trajectories. In BikePRINT, only trajectories chosen by clicking a road segment can be visualized together by means of a red color (Figure 4-12). These trajectories overlap each other and the user cannot know how many individual trajectories there are. Therefore, these cycling trips should be visualized differently. In addition, the origin, destination and direction of cycling movements on trajectories should be represented as well. The origin and the destination can be shown by differently colored point symbols and the direction can be represent by arrows of lines as introduced by Figure 10 so that other visual variables applied to lines can be used to represent other attributes (speed, intensity, etc.). When these trajectories could be distinguished by users, they may want to select one of these trajectories to get more detailed information. So it should be possible for users to select a single trajectory. They could select and highlight a single trajectory by clicking a unique part of this trajectory or by clicking its origin or destination. On the other hand, as introduced in section 2.5.2, small multiples are useful for users to compare different spatial patterns with different attributes or in different time, which could be added in BikePRINT.

3) Functionality (supporting data exploration)

As mentioned above, to achieve these possible improvements, some interactive functionality should be added to support users to explore the data. Firstly, users should be able to select the information they need or are interested in to have a detailed look. For example, when many trajectories are shown, they should be able to click on one of them to get more detailed information. And, in addition, the system should be able to record the users' operations so that they can undo their incorrect ones. For example,

users may want to focus more on information of a whole road than of a road segment. Therefore, it would be better if BikePRINT could aggregate road segments into a whole road. Finally, sometimes users need to calculate the average value of certain attributes and, therefore, a calculation functionality is needed as well.

The detailed design of these solutions will be described in Chapter 6.

4.8.　　Conclusion

In this chapter, the data of the case study were introduced elaborately from its background to characteristics. According to the dataset, I proposed some analytical questions that could be used by users to do their works. Then, visualizations of BikePRINT were described. There are some potential limitations that make that BikePRINT cannot support users to visualize and analyze trajectory data very well. So, the last part of this chapter consisted of a desk evaluation aimed at finding out limitations. Some analytical questions were used here to test whether I can find answers well with the help of BikePRINT or not. These limitations are only observed, and not tested by some user research techniques. Therefore, it is necessary to test them in the user research stage to find which limitations are needed to be improved. The next chapter is describing the whole implementation of the user research.

5.   USER RESEARCH IMPLEMENTATION

5.1.   Overview

In this research study, the user research experiment around targeted users, 'Fietstelweek' data and the web visualization environment BikePRINT was implemented. Several users were invited to participate. This chapter describes the whole procedure of user research implementation including three main components: (1) The questionnaires; (2) The actual user tests with task execution in Amsterdam and Breda; and (3) The Focus Group/interview in Amsterdam and Utrecht. In this Chapter, this procedure will be described from 3 steps: the test setup and preparation, the research execution and the results with analysis. The test set up preparation introduces the participants' selection, the tasks preparation, the questionnaires, the pilot testing and others. The part of execution introduces the whole procedure of testing with participants in Amsterdam and Breda. Then, the results and findings, such as collected data, users' needs and the analytical results, will be discussed. First of all, the objective of this user research will be introduced.

5.2.   Objective of the user research

This research aims to design an interactive visualization solution for analyzing spatio-temporal cycling trajectory data through conducting the user research with actual users. The 'Fietstelweek' data and its visualization solution BikePRINT are used as the case study. Through this user research, these information could be acquired:

1)   the profile(s) of the users of the 'Fietstelweek' data
2)   the spatio-temporal questions the users of the 'Fietstelweek' data want to find answers to
3)   the ways in which these users currently find the answers to the spatio-temporal questions they have
4)   the current limitations of the web application BikePRINT in providing answers to the spatio-temporal questions of the users
5)   other requirements of an improved BikePRINT application, i.e. the requirements of an interactive application for the visualization of 'Fietstelweek' data

The web application BikePRINT of 'Fietstelweek' data was designed to transfer cycling data to information and provide some visualizations to users. In order to have a better understanding of 'Fietstelweek' data and the application BikePRINT, meetings were held with some people who are involves in these projects. The first meeting was made with staffs of the leading company 'Keypoint' of the 'Fietstelweek' project in Enschede. And then, my supervisors and I went to Utrecht to meet with Mr. Joost de Kruijf, a member of the development team of BikePRINT. From these two meetings, two things were decided for the research. The first one is that some employees of the municipality who work in the domain of urban planning and analysis have used BikePRINT. The other one is that BikePRINT is not used as expected and this may be caused by usability issues. There may be different reasons of this: the visual representations of information are not desirable, the types of information provided is not users' needs, users spatio-temporal analytical questions cannot be answered well by the application, the features or the interactive methods of the application are not useful or not easy to use, and so on. In order to get a better insight into the variety of reasons that limit the use of the application, the intention was to conduct user research with groups of current users or potential users. As a result, the design of improved visualization solutions of cycling trajectory data can be developed. An additional advantage of doing this is

that the outcomes may be used to recommend improvements for the BikePRINT application. Therefore, it was decided to set up the user research which participants have to execute tasks with BikePRINT as part of the requirement analysis in this research, which is focusing on the design of an interactive visualization solution for analyzing spatio-temporal cycling trajectory data.

5.3.   The user research setup and preparation

5.3.1.   Plan of the the user research

As indicated above, the plan of conducting this user research with actual users to evaluate the usability of BikePRINT and analyze users' requirements includes three parts: (1) The questionnaire; (2) The actual user tests with task execution in Amsterdam and Breda; and (3) The Focus Group/interview in Amsterdam and Utrecht. The part 1 were made via the Internet and the part 2 and 3 took place with participants face to face.

In Part-1, the purpose of sending questionnaires is to acquire users' backgrounds, their purposes experience of using 'Fietstelweek' data and BikePRINT, their interested analytical questions and their wishes, etc. The questionnaires will be firstly sent to participants who will doing task so that the tasks used in Part 2 could be made on the basis of feedback. Then it will be sent to other users of 'Fietstelweek' data and BikePRINT which aims to acquire information mentioned above from more users.

Then, in Part-2, participants will be invited to do some analytical tasks under the environment of web application BikePRINT one by one in Amsterdam and Breda. The contents of tasks are made on the basis of their interested analytical questions that I provided in the 'questionnaire' or they proposed. When one user is doing tasks, she/he should think aloud to tell me what are in his/her mind at that moment. I will observe the whole procedure with using the video,audio and screen recorder 'Camtasia' (URL 4) to record it so that their activities can be captured and stored as video files. By these means, the application could be tested and limitations could be found out. In this step, limitations that I found will be verified as well. Also, only participants in Amsterdam will participant in Part-3.

In Part-3, focus group discussion/interview will be conducted in two places. The first discussion will be made with two participants in Amsterdam after they finished tasks. In the discuss, we will first discuss found limitations with them and then present some possible improved visualization solutions to them according to found limitations by the help of PPT slides. They could evaluate these solutions and give their comments and wishes. Then, the second discussion will be made with Mr. Joost de Kruijf in Utrecht who is one of developers of BikePRINT. In the meeting, I also present solutions to him and he shared some their designed solutions.

After the implementation, the results were analyzed. By these approaches, the limitations of BikePRINT and user requirements could be Figured out to be used for the conceptual design of an improved BikePRINT application and general recommendations for the design of an interactive visualization solution for analyzing spatio-temporal cycling trajectory data. Figure 5-1 shows the flow of this user research with participants.
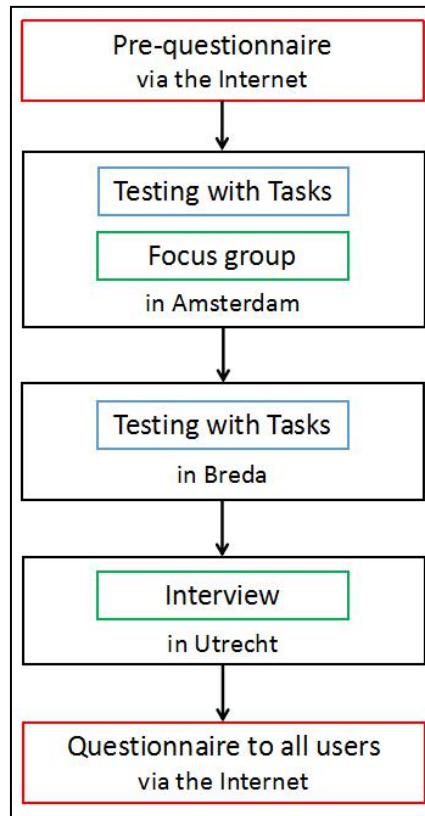
Figure 5-1 The flow of the user research. Two elements in the red rectangle belong to Part-1. Two elements in the blue rectangle belong to Part-2. And two elements in the green rectangle belong to Part-3.

5.3.2. Testing participants and groups

The participants of this user research were invited by the help of Mr. Joost de Kruijf, a member of the development team of BikePRINT. Of these participants, 2 are employees of the municipality of Amsterdam, 2 are researchers of NHTV Breda, and another one is the employee of the Province of Noord Brabant. Their specific profiles are collected by the assigned questionnaire. The reason why we invite them is that these people actually involve in transportation/cycling-related works so that they have practical needs for their analysis of urban cycling. Because of the arrangement and users' schedule, they are met as two groups in Amsterdam and Breda respectively. In Amsterdam, there are 2 employees of Amsterdam municipality participated in the user research. And other 3 participated in Breda. Because of a time issue, we could not invite more participants but the limited number of participants still brought many findings for us.

5.3.3. Questionnaire

The questionnaire (Appendix 1) were sent to users. As purposes of the questionnaire mentioned above, there are three components in the questionnaire to meet these purposes. The first one is designed to collect participants backgrounds and experience in using maps, 'Fietstelweek' data and BikePRINT. Then, in the second part, all existing BikePRINT visualizations are provided to them. The purpose of setting this part is to find which visualizations in BikePRINT are users' interests or are commonly used by them. Users also are asked to write down comments about these visualizations which, for example, include the purposes of using these visualizations, the visualization's limitations they found, analytical questions were not answered well by visualizations, their wishes of these visualizations and etc. So from the results, we can know users' interested visualizations in BikePRINT and focus on improving them. And in the last part of questionnaire, all proposed analytical questions are provided to them to choose and they are asked to

provide their own interesting analytical questions. From the feedback, we can know: which visualizations in BikePRINT and analytical questions are users' interests or needs.

5.3.4. Preparation of documents and devices

Before going to implement the task testing and focus group discussion, some documents should be prepared to guide the procedure which are: hard copies of instruction and task forms (Appendix 2), hard copy of the scenario (Appendix 3) and the plan of focus group (Appendix 4). The instruction is a introduction of the testing for users. They can understand the plan quickly and follow it by reading the instruction. The task form will be provided to users to do the tasks in it by using BikePRINT. The scenario is like a detailed schedule which guides me to conduct the user research. Besides, some other devices are also needed to prepare. For example, the laptop with software 'Camtasia' (for recording users' activities when using BikePRINT).

5.3.5. Conclusion

This section introduced the preparation stage of the user research. It is necessary to conduct it following a well - considered plan. The next section is to introduce the whole procedure of implementing the user research.

5.4.   Implementation of the user research

According to users' time, I made the user research with two participants in Amsterdam at first and then with three in Breda as planned. Two participants in Amsterdam firstly received questionnaire and were mat to do task testing and focus group discussion. And then for three participants in Breda, I only made Part-1 and 2 with them. In following sections, I would firstly introduce the result of questionnaire with these all 5 participants. And on the basis of these results, some tasks were formulated which will be described as well. Then I will introduce two procedures in Amsterdam and Breda.

5.4.1. Questionnaire assignment and feedback

Questionnaires were sent to users at two times: before and after the implementation of part 2 and 3. In this section, only results of first time will be described because the contents of the task testing were made on the basis of these results. At the end of this chapter, the results of second questionnaire assignment will be summarized because it was sent after the implementation of part 2 and 3. When I received email addresses of people who will participate in the user research, I sent the questionnaire to them via email. Results of three parts of questionnaire will be introduced below.

Part 1: Table 1 shows the summary of the feedback of first part. As we can see from the Table 5-1, all participants' works both are related to urban transportation and cycling analysis. They use cycling data to analyze cycling use and bottlenecks of the cycling infrastructures. Identifying the fast and popular routes in urban areas are focused by them as well. Other important information they would like to know more are about cycling trips and cycling behaviors in which cycling speed and directions of cycling trips are their interests. Therefore, they could have opportunities to improve cycling infrastructures that bring better experience for cyclists if BikePRINT could support their analysis very well.

Of these participants, 1 and 2 who works for Amsterdam municipality both focus more within the area of Amsterdam municipality. They have more experience in using BikePRINT and use maps very often. For Participant 3,4 and 5 focus cycling information of more than one municipalities but they have less

frequency in using maps (monthly or annually) and less experience in using BikePRINT (never or few times).

Table 5-1 The summary of the feedback of first part. As we can see, They focus more on cycling and infrastructures use, the fast and popular routes, speed and directions of cycling trips and cycling behaviors.

Participant No.	Works	Experience/purposes of using 'Fietstelweek' data and other cycling data	Focused research area	Experience in using maps	Experience in using BikePRINT
1	Transportation planning	Analyzing cycling speed, direction and bottlenecks of infrastructure	Within a municipality	Daily	Several times per month
2	Transportation planning	Analysis of cycling use	Within a municipality	Several times per week	Several times per week
3	Cycling program of province of Noord-Brabant	Getting better cycling data of Noord-Brabant, identifying fast cycling routes and acquiring needs of improving infrastructures	More than one municipalities	Several times per month	Never
4	Cycling project in NHTV	Finding different routes of new infrastructures and knowing more about cycling trips and behaviors (sometimes use it)	More municipalities in Noord-Brabant	Several times per year	Only once/twice
5	Mobility management Urban intelligence	Using data for knowing amount of cyclists	A city and its surroundings	Several times per month	Few times per month
6	Cycling researcher and cycle project initiator	Counting data (cross section and traffic lights) to determine in- and decrease of cycling volumes. Evaluation of infrastructural improvements	A city and its surroundings and More than one municipalities	Several times per month	Several times per month

Part 2: The second part of the questionnaire is about all visualizations in BikePRINT. Users could choose which visualizations they have used before, they may want to use or they are not interested in. Then, they wrote down comments about these visualizations. Because the participant 3 (P-3) never used BikePRINT, he only could show his interests of visualizations by choosing 'I may want to use it' and give comments. From the feedback, there are some conclusions are made. Firstly, 'Intensity', 'Route', 'Absolute speed' and 'Delay' map are commonly used by them or more users wish to use them. So we introduce comments about these visualizations in detail.

For 'Intensity' map, all users (except participant 3) have used it and felt that it is not easy to extract proper information from the visualization. P-1 wishes to be able to make selections of road segments. P-2 would like to select routes for average intensity And P-4 thought that it gives a rough idea of the busy parts/streets and information is not specific and colors are not good chosen which can be improved. As the result, I could conclude that they wish to be able to select the road segments and some of them thought that color visual variables used in the map are not very desirable. Then, for 'Route' map, almost all users have used it. P-1 wishes to know the percentages of the cyclists are on the sub-branches. P-2 would like to know amounts of cycling trips of selected road link and he would like to select the road links as origin/destination of cycling routes. P-3 wants to use this more to identify the best route for fast cycling routes. And P-4 needs more information about the cycling trips. Therefore, some thought that it is needed to provide more information of cycling trips such as amounts of cycling trips, origins/destinations and etc. And for P-3's wish, cycling trips should be analyzed with the combination of speed information. As to visualization 'Absolute speed', P-1 and P-2 need the possibility to calculate the average speed on a route. P-4 thought that visual variables used in this visualization are not clear. And the meaning of different colors of road segments needs improvements. For 'Delay' map, P-2 and P-4 commented that the better definition and explanation of delay points should be provided. In addition, more detailed information of these points are needed by them.

From the feedback, we could know that users use/are more interested in 'Intensity', 'Route', 'Absolute speed' and 'Delay' map mostly. Therefore, these four visualizations are emphasized and analytical questions will be made on the basis of these four visualizations.

Part 3: The last part is about analytical questions which users are interested. As only 5 participants gave feedback before the implementation of the task testing and they have various focuses, we couldn't conclude which analytical questions are really important among all analysts. So only some popular questions of them are extract to formulate analytical tasks with the consideration of interested visualizations. There are some analytical questions that most of them focus:
1) What is the average number of cyclists on road segment X per day?
2) Where do cyclist ride bikes in city X on day Y?
3) There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?
4) Where are the destinations of cyclists from the central station of city X during the morning rush time of day Y?
5) What is the trend of movements between city X to city Y during the whole period?
6) What are the spatial-temporal patterns of trajectories from city X to city Y in the morning rush time and night rush time of working days? What is the difference?

On the other hand, these are questions that only one or no participant focused so we could do not considered them very much:
1) Where is the longest cycling trip in city X on day Y? Where are its origin and destination?
2) Where is the earliest/latest cycling trip in city X on day Y?
3) Where is the cycling trip that lasts for the longest time in city X on day Y?
4) On which day did most cyclists depart from the central station in city X?
5) What are the spatial-temporal patterns of cycling trajectories from the east to west part in city X on day Y?
6) What is the spatial-temporal distribution of cycling trajectories in the city which have the highest average speeds on day Y?
7) What are the spatial-temporal patterns of cycling trajectories in the city X during the whole period?

5.4.2. Analytical tasks

Through the consideration of interested analytical questions, 7 main tasks were made for different purposes which is consist of several small tasks/questions (APPENDIX 2). It is known that participants are interested in using 'intensity', 'routes', 'absolute speed' and 'delay' maps more. Therefore, these tasks are made on the basis of these maps. Then, for interested analytical questions above, these tasks were made according to the theories in the section 4.4.2. The main purpose of tasks are described below. Task 1 is related to the visualization 'Intensity' and meant to test the selection of provided information (e.g. finding busiest road segments from all). Task 2 is more about cycling trips. Participants are asked to find all trips between two cities and their origins, destinations and directions. Task 3 is also related to cycling trips more and their relations with the given place (the central station). Moreover, the comparison functionality of BikePRINT is also tested in this task. Then, in task 4, temporal information of cycling trips are tested. And task 5 is meant to test how people analyze cycling routes and speed together. Task 6 is related to visualization 'delay'. Users are asked to find bottlenecks in the cycling infrastructures and make selection of them. Then, the last one task is asking users to select a cycling trip to get more detailed

DESIGN OF AN INTERACTIVE VISUALIZATION SOLUTION FOR ANALYZING SPATIO-TEMPORAL CYCLING TRAJECTORY DATA: A CASE-STUDY OF THE
VISUALIZATION OF FIETSTELWEEK DATA WITH BIKEPRINT

_____

information of it. For example, the attributes of all road segments that constitute a trip are asked to find out. In every task, there are some questions to ask users how do they think the application supports them to complete tasks.

5.4.3.   Pilot testing

The purpose of conducting the pilot test is to check how long the whole test costs and to find out problems in tasks procedure including language errors, logical problems and etc. So improvements of tasks could be made before the real experiment implementation. On 23th February 2017, I invited my first supervisor Dr. C.P.J.M. van Elzakker to attend the test because he could give many professional suggestions. During the pilot testing, the real test scenario was executed. And after that, some corrections of tasks were made as the result of the pilot test. And the approximate testing time was estimated which is 40 minutes.

5.4.4.   Testing procedure in timeline

As there are two groups of participants in Amsterdam and Breda respectively, I went to these two places to meet with them. Two participants are in Amsterdam (No.1 and 2) and three are in Breda (No.3, 4 and 5). On 28th February 2017 in Amsterdam, I first invited them to do tasks one by one. And then, a focus group was done with both participants. At beginning of the focus group, we discussed some results and findings which I observed in the testing. Then, I presented some visualization concepts of trajectory data and improved solutions according to the limitations I found beforehand to them. They shared their ideas and needs about them and gave comments. As these contents and comments are related to visualization solutions, I write dome them after every corresponding solution in the next chapter. Then, on 9th March 2017, I went to Breda to implement the second user research with the other three participants. The testing environment is shown as Figure 5-2. However, this time I only invited them to do tasks and there was no Focus Group. Then, on 16th March, I had a meeting with Mr.Kruijf in Utrecht to discuss solutions and share ideas, which will be also discussed in the next chapter because contents of the meeting are mostly related to improved solutions. Therefore, in this next section, all procedures of task testing part of the user research will be described.
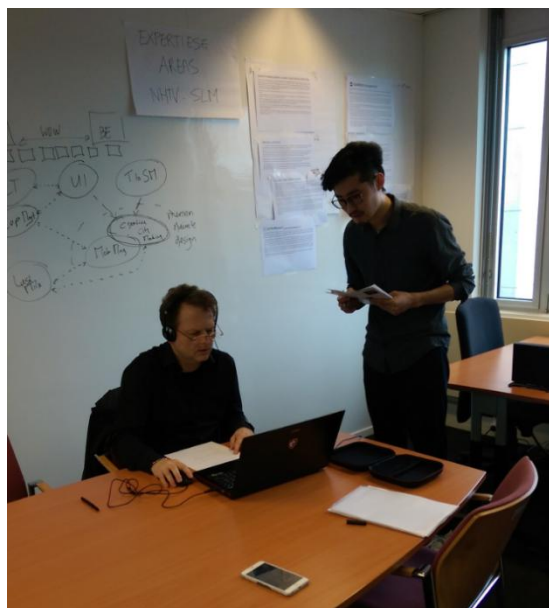
Figure 5-2 Testing environment in Breda

5.5. Results of the task testing implementation

In the previous sections, I introduced and discussed the overall set-up works of the user research. This section deals with the results gotten from the research implementation. In this section, we mainly focus on the performance of 5 participants who have completed tasks. The visualization and usability issues of BikePRINT can be extracted from the recordings of their activities during the testing.

5.5.1. Test Results of the task form

In the task form, after every sub-task/question in these 7 main tasks, I set the question 'can you find it/them?'. Users could choose yes or no and then give comments or reasons why they think they cannot find the answer. Table 5-2 shows the results of these choices. Those comments will be described with their performance in the next section. In this Table, 'Y' means the user can find the answer, the 'YC' means the user thought he/she finds it but not sure and 'N' means the user cannot find it.

Table 5-2 The results of questions in task form

Participant No.	The answers of the question 'Can you find it/them?' of each task																	
	1-a	1-e	2-a	2-c	2-e	3-a	3-b	3-c	3-e	4-a	5-a	6-a	6-c	6-d	7-a	7-b	7-c	7-d
1	N	YC	N	N	YC	Y	N	N	N	N	Y	N	YC	N	N	N	Y	N
2	Y	N	N	Y	Y	Y	Y	N	N	N	Y	Y	Y	N	N	N	N	N
3	Y	N	N	YC	Y	Y	Y	N	N	N	Y	Y	Y	N	N	N	N	N
4	N	Y	N	N	YC	Y	N	N	N	N	Y	N	N	N	N	N	N	N
5	Y	Y	N	N	Y	Y	Y	N	N	N	Y	Y	Y	N	N	N	N	Y

As we can see, for most questions, participants cannot find answers to with the help of BikePRINT very well, especially for task 2-a, 2-c, 3-c, 3-e, 4-a, 6-d, 7-a, 7-b, 7-c and 7-d. For these tasks, they also gave some reasons that lead to some answers cannot be found from their perspectives. For task 2-a, all users cannot find these trips and origins/destinations because there are no visualization of them on the map. For, task 2-c, three participants cannot find related temporal information and two thought they found that but the answer is wrong. For task 3-c, users can some trips that passed the central station but not all and both of them didn't find origins/destinations. For task 3-e, they tried to compare the patterns but didn't find the helpful functionality. For task 4-a, same as the result of task 2-c, they didn't find visualizations of temporal information of trips. For task 6-d, all of them cannot only show bottlenecks because they cannot find functionality to filter others out. For task 7-a, all participants cannot select one single cyclist trip. The participant 2 wrote down: 'I cannot select the single cycling trip from these trips so I cannot know its detailed information such as the direction'. So he gave 'No' to this questions. This means users cannot select any single cycling trip in BikePRINT if they would like to analyzed details of it which is a lacking functionality of BikePRINT. so for the following task 7-b, 7-c and 7-d, they cannot complete. These information will be analyzed in following sections along with their other recordings.

In the task form, I set 9 questions 'How well did the application assist you in answering this question?' to acquire users' satisfaction. And Table 5-3 shows how users think BikePRINT helps them to complete tasks in which very good, Good, Average, Poor, Very poor are represented by 'VG', 'G', 'A', 'P' and 'VP' respectively. From the result, we can find that to 1-f, 2-d, 3-f and 4-c, they felt BikePRINT cannot assist them very well. For 1-f, although most of participants thought they can find answers but they commented that they only can guess the answers. Therefore, they thought BikePRINT cannot help them a lot. For 2-d, because most of users cannot find answers to 2-a and 2-c, so they gave negative comments to this

_____

question. As to 3-f, most of users cannot find answers to 3-c and 3-e and they cannot fidn some functionality in BikePRINT to help them find answers, so they thought BikePRINT provides (very) poor helps. And for 4-c, the answers of 4-a are both 'no'. Both of them cannot find related information in BikePRINT.

However, for some tasks, users attitudes are various for three reasons: (1) Users have different understandings of tasks and visualizations because they are not experts in cartography domain; (2) Some answers that they thought are right actually are wrong so they gave different remarks to others; (3) Because I didn't give an unified standard of these choices so that they may follow their standards. For example, there is a task 'Please find the 10 busiest road segments in Eindhoven'. Actually, in BikePRINT, people cannot do it because they only can guess where are these 10 road segments but cannot find segments with specific numbers. So the right answer should be 'No, I cannot find.'. However, two participants chose 'Yes'.

Table 5-3 Users' satisfaction of tasks

Participant No.	Users' satisfactions of tasks								
	1-c	1-f	2-d	2-f	3-f	4-c	5-d	6-e	7-e
1	P	P	P	A	P	VP	P	P	P
2	A	P	VP	G	P	VP	A	A	VP
3	VG	VP	G	VG	VP	A	A	G	A
4	A	G	P	P	VP	P	P	VP	VP
5	A	P	VP	VG	VP	VP	A	G	A

As only five participants attended the testing, this analysis of the results are relatively less useful but it still reflects some limitations of BikePRINT. And further, this should be analyzed together with all participants' performance which are contained in the video/audio and screen recordings. Analyzing these data attentively could reveal some usability and visualization limitations of BikePRINT.

5.5.2.   Processing results of participants' performances

The first step of analyzing these recordings is to transcribe and interpret these results to texts. By this way, the information could be more easily extracted and understood. All five video files of five participants were looked through with the help of the website oTranscribe (URL 9). Because of the time issue, not all activities or voices were interpreted but only the relative information of limitations. Figure 5-3 shows the working page of website. The result was a document with 2468 words for these five participants. The content of the document are mainly the moments when they feel confused, they go to the wrong visualization page, they hesitate, their failed tries, some key words they said and etc. Different codes were to group these contents and identify the hidden factors that lead to these results. In this work, the qualitative research software Atlas.ti (URL 5) was used. Before using the software, a code system was defined according to this research. Table 5-4 is used code system in which they represent some key activities and performances of users. Then, I looked the document and inserted codes to the corresponding positions in software Atlas.ti, as shown by the example of Figure 5-4. Then, the key problems all participants met or their own problems of using BikePRINT could be extracted and summarized, which will be analyzed in the next section.
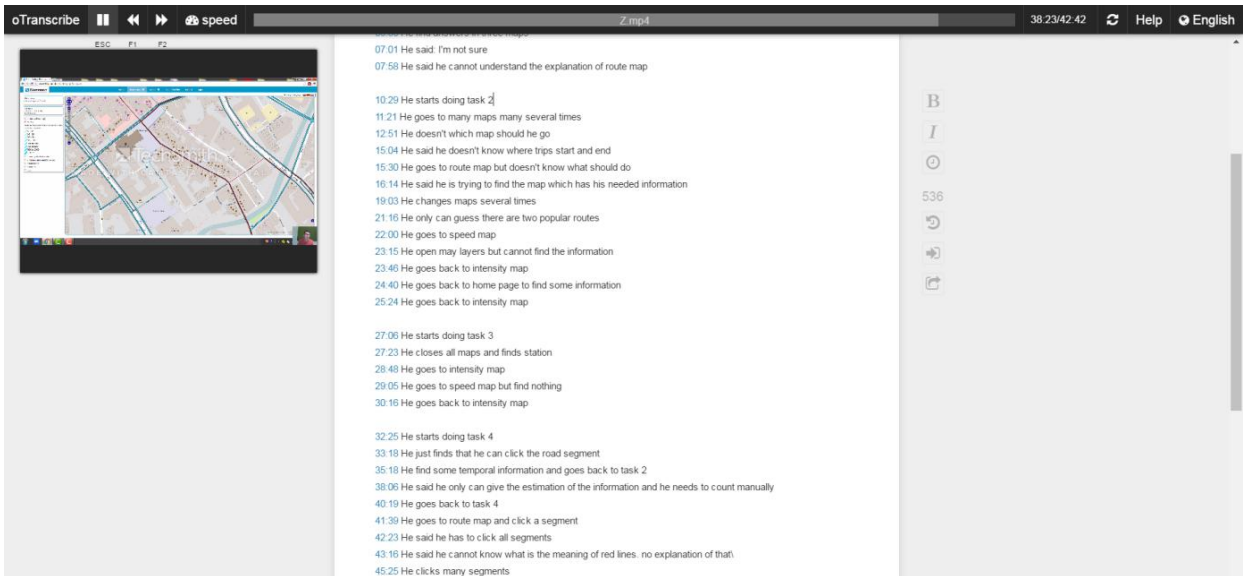
Figure 5-3 The working page of website oTranscribe.

Table 5-4 The code system used in this research

Codes	
Multi-clickings	Confusion
Verbal comments	Hesitation
Change maps	Errors
Visualization issue	Access data issue
Failed try	Ask for help
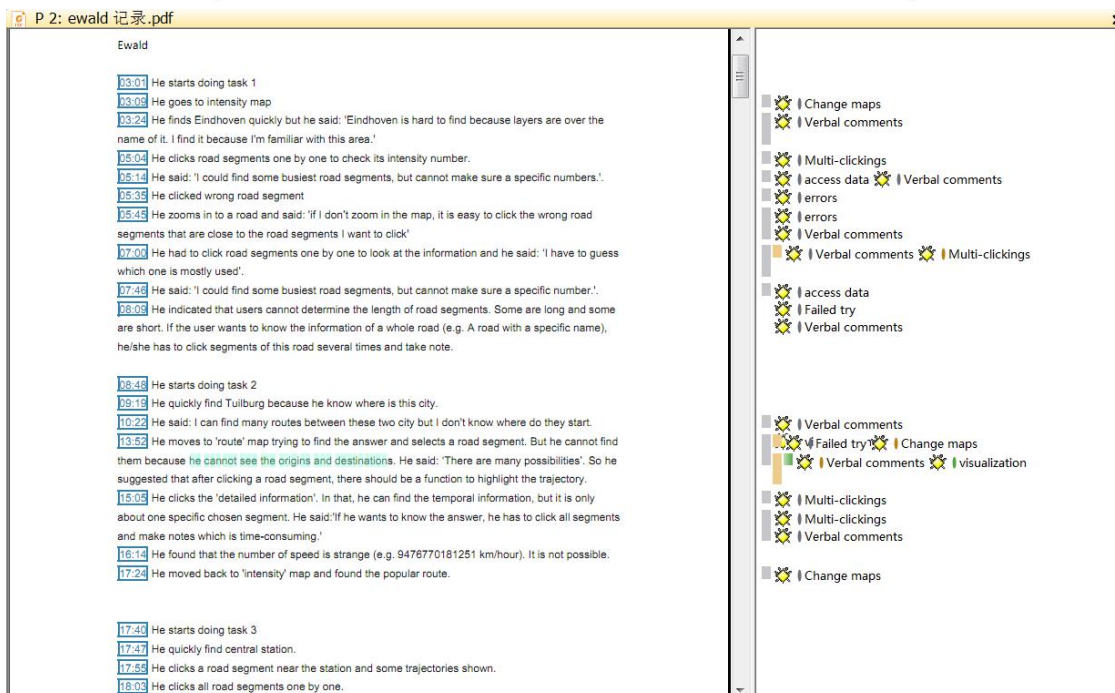
Figure 5-4 The example of a working page of coding in Atlas.ti

5.6. Analysis of results

In this section, at first, all five participants' key activities and problems they met will be described respectively. And then, the general During the testing, not only these five participants met same problems but also they had their own different performances and problems when they use BikePRINT to do tasks. In this section, their performances will be described one by one. And then, the similarity and differences will be summarized in the next section. The performance of the participants which extracted by the help of oTranscribe and Atlas.ti are introduced as follow:

5.6.1. Time cost of each task of participants

The Table 5-5 show the time cost of each task of each participant, average and sum time. As we can see, the participant 3 spent most time in doing tasks because he never used it before. In addition, most participants spend more time in task 1, 2 and 6.

Table 5-5 Time cost of each task

Participant No.	Time cost							
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Sum
1	5'47"	8'52"	2'34"	48"	1'8"	5'33"	1'9"	25'51"
2	6'36"	11'24"	6'26"	57"	2'29"	6'44"	3'51"	38'27"
3	10'20"	16'37"	5'19"	15'4"	3'32"	9'12"	8'22"	1'8'6"
4	6'33"	8'29"	6'39	4'6"	4'9"	8'33"	4'35"	43'4"
5	3'40"	4'28"	3'21"	55"	1'58"	2'19"	4'42	21'23"
Average Time	6'31"	9'39"	4'33"	4'44"	2'30"	6'28"	4'32"	39'22"

5.6.2. Results of participants' performances

1)   Participant 1 have used BikePRINT several times before so he is familiar with contents in BikePRINT. In task 1 (finding busiest road segments), he can quickly go to the visualization 'intensity' to find busy road segments. Because he have used BikePRINT to do the similar task (find road segments with high intensity and find amount of them), he knew that he cannot do it in BikePRINT and said: 'You can see some busy road segments on the map by the legend but you cannot make sure which is busiest unless you click all road segments one by one, which is time-consuming.'. In his testing, he clicked some segments and guess that a segment with value of 556 is the busiest one. And he also met another problem that sometimes he would click wrong segment. And in task 2 (finding cycling trips between two cities), he cannot find Tilburg easily because the name is covered by the layer. And he went to some visualizations to find cycling trips (trajectories), which was coded by 'Change maps' several times. In order to find cycling trips, he also clicked road segments many times but finally he said: 'I cannot find the number and direction'. There are many code 'Multi-clickings' were marked here. In task 3 (finding cycling trips passed the station), he also clicked adjacent streets of the station many times and find some trips that passed the station but not all. Similarly, origin and destination are unknown. Then in task 4 (finding temporal information of cycling trips), He had to manually click every road segment and count the number by himself, which is not possible for him so he said he cannot do it. As for task 5 (analyzing cycling routes and speed), he can have the overview of segments with speed value of more than 20km/h but cannot know the exact number. And in task 6 (finding bottlenecks), he quickly went to 'delay' map to find answer. He

thought that bottlenecks should be considered by the combination of a lot of delay and a lot of cyclists (color and size of points). He also said there are some clutters on this map. In task 7 (finding detailed information of a cycling trip), he can select road segments to show some trajectories but cannot select a single one from them. He indicated that some direction information could be found from the base map. Also, code 'Multi-clickings' was inserted here many times. And he cannot distinguish road segments.

2) Participant 2 uses BikePRINT commonly so he is very familiar with BikePRINT. In task 1 (finding busiest road segments), he said 'I could find some busiest road segments, but cannot make sure a specific numbers.', which is same with the participant 1's response. And he also indicated that he cannot distinguish road segments very well. For task 2 (finding cycling trips between two cities), he could find some routes but he cannot find origins and destinations. Also, temporal information of cycling trips are unknown. When he tried to find routes between these two cities, he said: 'There are many possibilities'. In task 3 (finding cycling trips passed the station), he clicked streets many times, which is coded by 'Multi-clickings'. Towards to these shown trajectories, he said 'there are three types of trajectories shown together: (1) passing this segment, (2) starting at this segment and (3) ending at this segment. I cannot distinguish them.', which is coded by 'visualization issue' because it was caused by undesirable visualization results. In task 6 (finding bottlenecks), he needs a better definition of 'bottlenecks'. And he was confused by visual variables of delay points. He also thought that bottlenecks should be considered both with amount and time. For task 7 (finding detailed information of a cycling trip), he knew he cannot select the single trajectory by his experience. When he went to 'speed' map, the chosen trajectories are disappeared. Because he is familiar with the BikePRINT, he can quickly go the the right page to find answers. And he can understand the visualizations of BikePRINT and tasks very well.

3) Participant 3 never used BikePRINT and he spent most time to complete tasks(more than 1 hour). He has same problems mentioned above when he used BikePRINT and he also has other problems. In the testing, he cannot understand some information expressed by visualizations and some explanations of them (e.g. In 'intensity' map, 'routes' map and 'speed' map). Therefore, he was coded by 'Confusion' several times. And then, he cannot precisely find the right visualization for the task so he visited many visualization pages to find the answer or stopped to consider where should he go, which was very time-consuming and coded by 'Change maps' and 'Hesitation' many times. He said 'I'm trying to find the map which has my needed information' two times. Almost in every task, he did change maps several times. He said that when he waits for the application to react, he doesn't know it's status (no signal to show that). So he doesn't know should he move or not.

4) Participant 4 is only one female participant and only used BikePRINT once or twice. In task 1 (finding busiest road segments), she clicked road segments to get values. She said she know where are segments with high value but doesn't know how to select them. In task 2 (finding cycling trips between two cities), she tried to draw a circle to select some road segments but failed. Also in task 3 (finding cycling trips passed the station), she tried using 'Ctrl' to select more segments but also failed. An important information she said is that she is not sure whether BikePRINT doesn't provide the functionality or she cannot find it. She tried clicking legend to select segments but failed again. During the testing, she got the code 'Failed try' mostly. She also have some confusion about bottlenecks.

5) Participant 5 only spent 20 min to complete tasks. When doing task 2 (finding cycling trips between two cities), he clicked one segment in Eindhoven and some passed cycling trips were shown. And then he clicked another segment in Tilburg. But the shown trips of Eindhoven were disappeared. He said that it is not he wants to see because he thought the first point he clicked is origin and the second is destination. He also visited many visualizations to find answer. Similarly, he cannot find origins/destinations and directions of cycling trips. When he went to speed map from route map, he found that the chosen trips were disappeared so he said: 'I may need a functionality that I can use to go back'.

5.6.3.   Summary of their performances

Through analyzing their performances, I found that there are some common that most of participants have met. These problems lead to that they cannot find answers to the questions or they cannot make use of BikePRINT very well. Then I summarized these problems as follows with the their revealed limitations:

1) Users can have overview of some road segments with values. In other words, they can see the road segments that they are look for (e.g. busy road segments). However, they cannot select some of them to get detailed information or even obtain the number of them. If the user wants to know some detailed information of segments or cycling routes, they have to manually click all of them one by one many times and take notes, which is very time-consuming to them. So two of participants said that: 'I know I can do it but I won't because it will cost me too much times'.
- **Limitation**: BikePRINT was designed for planners to analyze the data of cycling and infrastructures usage. It is not enough for them to only see these information. The amount of road segments with a particular value is useful for them to do some detailed analysis of road segments. And then, users should be able to select some road segments instead of only can select one. Therefore, the reason was concluded that users cannot have the functionality of filtering road segments their needed attributes (e.g. Values or times).

2) It is not easy for people those who are not familiar with the targeted area to find or locate some cities when information layers (intensity, route, speed...) are switched-on because names of cities are on the base map that could be covered by layers. And some of them also thought that the base map of is too colorful and close to some colors of layers that may effect users' understanding.
- **Limitation**: The colorful base map sometimes makes people cannot quickly focus on meaningful information.

3) Sometimes, the users want to know the information of a whole road (e.g. With a specific name) in stead of road segments. So if they want to do that, they have to click all road segments that constitute this road and take notes. In BikePRINT, when the user clicks a segment, this segment is not highlighted. So sometimes they cannot distinguish road segments (some are long, some are short.) and they cannot know which segment they are selecting. And they cannot know where exactly are the chosen road segment.
- **Limitation**: When the user clicks a road segment but it is not highlighted, he/she cannot know where exactly is this road segment. Only selecting one segment once cannot meet their needs very well. And they even cannot know how many road segments constitute a road so they cannot analyze the information of this road.

4) The only way of show some trajectories in BikePRINT is to click a road segment and see passed trajectories. Users have no other ways to select them. When they click a road segment to show all cycling trips that passed this segment, they cannot distinguish them (they overlap each other) and

they cannot select a single trajectory/trip from them. They also cannot know the amount of these trips. Of these cycling trips, the direction, origin and destination of the them cannot be visualized.

- **Limitation**: Is it not enough to select cycling trips only by clicking a road segments which will generate limited cycling trips. Some people also don't know what is the meaning of clicking road segments to show passed trips. How do this way support analysis. So more ways of selecting cycling trips which have common characteristics (e.g. Same origin/destination) could be much useful for analysts. In addition, good visualizations of these trajectories are also necessary.

5) For some participants who are not familiar with the BikePRINT, they cannot understand what are shown in BikePRINT very well. For example, some of them cannot understand red lines in the 'route' map, points in the 'delay' map. They have to guess what are meanings of these information by their experience but cannot get clear explanations or definitions from BikePRINT.

- **Limitation**: BikePRINT should improve explanation of what it shows, especially of 'delay' map and 'route' map. So, users could precisely go to the right visualization to analyze the data.

6) In the task 6, users were asked to find bottlenecks. However, they have different understandings of bottlenecks. And most of them don't know where should go to find the information so they visited many visualizations to find answers. And in 'delay' map, some of them thought that bottlenecks should be considered with the combination of 'delay time' and 'delay amount' of each delay point. However, they cannot extract these points in BikePRINT. And some of them also felt confused about size visual variables of these points.

- **Limitation**: Same as the previous one, they cannot directly go to the right visualization because they are not clear of information provided by visualizations. Analyzing bottlenecks by combining two attributes ('delay time' and 'delay amount') needs the functionality that could extract or filter these attributes and visualize the result.

7) Users' operations in BikePRINT cannot be recorded so that they cannot 'undo' or go back to previous clicking/activities. For example, the participant 5 clicked a road segment in 'route' map and then he wanted to find speed information of chosen cycling trips. When he went the the map 'speed', he tried to chick another road segment to know detailed speed information. However, he cannot find the previous chosen cycling trips.

- **Limitation**: besides to the issue of the users' operations, it also reflects another limitation that users cannot analyze different data in one visualization page, or in other words, BikePRINT cannot visualize different data together (e.g. Visualize trajectories together with intensity/speed information of road segments).

8) Users only can find temporal information of a chosen road segment in BikePRINT (e.g. Number of cyclists who passed this road segment at different time). They cannot find patterns of road segments or cycling trips in different particular time intervals

- **Limitation**: It is not enough for them to analyze and answer questions only with the help of these limited temporal information. They also cannot count the amount of them (amount of cyclists). They lack a functionality of filtering elements (e.g. road segments or cycling trips) by time.

9) There are some clutters in maps:
   a) Without enough zooming-in, delay points generate some clusters which also happens in visualization 'difference measured/shortest';

b) Without enough zooming-in, some road segments overlap so that users always click wrong segment.

- **Limitation**: It would also effect users' reading and selection of wanted information.

5.6.4. Conclusion

These problems and limitations lead that users cannot use BikePRINT to analyze the data and complete tasks very well. Of these problems and limitations, most are similar with limitations I found before which were described in section 4.6. Firstly, BikePRINT cannot provide functionality of filtering values and times for users to extract their needed information very well. Users cannot select or choose their needed information. For example, they cannot extract all road segments with that intensity is more than 500. So users only can access all information that BikePRINT provides which may be not their needs. Then, for accessed data, BikePRINT cannot provides the good visualizations of them. The visual variables used to show attributes of road segments are not suiTable. And for cycling trajectories, their origins/destinations and directions cannot be shown. When there are many trajectories on the map, users cannot distinguish them. And there are some new findings, such as different performances of 'professional' users and new users. New users cannot understand the meanings or elements of some maps very well (e.g. the points in 'delay' map). Moreover, some wishes/requirements of users were obtained during the testing which would be considered as well.

5.7. Results of questionnaire to other users

Till the time of submission of thesis, I have received feedback from 11 users.

5.7.1. Their backgrounds and experience

From the feedback, 9 of them are working in transportation/cycling domain.

And they use 'Fietstelweek' data mainly for: (1) investigating the use of the road network and cycling routes (5 participants); (2) getting the amount/intensity of cyclists on the road network (4 participants). One focuses on the origins/destinations of cycling trips and another is interested in the relation between cycling routes and the interesting point (central station).

Of these users, 5 focus on more then one cities, 4 focus on a city and surrounding areas, 3 are interested in the urban area of a city and only one considers about the whole country.

As to experience in using maps, 5 of them use map daily, two use map several times per week, two use map several times per month, one uses maps several times per year and another one never used maps.

Then, their experience in using BikePRINT were also acquired. 6 persons only used it one or twice, 3 use it several times per week and 2 never used it.

5.7.2. Users' comments about visualizations of BikePRINT

In this part, only 10 participants give comments about visualizations. The comments will be described by each visualization and then the summary will be given.

- For 'intensity' map, 3 participants showed interests on it and 7 have used it. 4 participants thought it is useful and could provide their needed information. They used it for making policy of cycling routes. But two of them thought it is better if its visualization result could be improved.

- For 'route' map, 4 participants showed interests on it and 6 have used it. 3 of them thought it is useful for their works. And 4 of them need direction information of cycling trips. 2 wish to be able to select cycling trips by more ways. And 1 is interested to know origins and destinations of cycling trips.

- For 'Intensity shortest route' map, 4 of them showed interests on it and 4 have used it. 3 people thought it is difficult to understand and interpret the information of it. They need more definition of this map. And other 2 people thought it is good information.

- For 'Difference measured/shortest' map, 3 of them showed interests on it and 4 have used it. Others are not interested in it. Four of them thought it is difficult to understand what are expressed so that they needs more clear definition of it. 2 of them thought it could be used to analyze cycling routes.

- For 'Detour factor' map, 7 of them showed interests on it and 2 have used it. 3 of them thought that it is difficult to understand what are expressed by this map. 2 of them thought it is useful to find missing links in the network so that they can improved it.

- For 'Absolute speed' map, 4 of them showed interests on it and 4 have used it. 3 of them thought it is an interesting map and useful. But by one's experience, it is less useful outside the city area's. They commented that is could be more helpful if there is a functionality that people could use to calculate the average speed of a route or between 2 points. Moreover, one indicated that it is better to be combined with 'delay' map.

- For 'Relative speed' map, 4 people are not interested in and 3 showed interests on it. 2 of them have used it. They gave less comments about this map. 2 of them thought it is difficult to understand the information of it and one thought it needs better visualization.

- For 'Delay time' map, 3 have used it, 4 of them showed interests on it and 3 people are not interested in it. Most of them needs more detail information and sum information of delay points. They also indicated that the better definition is needed so that users can focus more on the most problematic points

- For 'Isochrones' map, 6 have used it, 3 of them showed interests on it. Two of them thought it is useful. They use it mainly for making policy about and compare bike routes. It is also helpful for them to find missing links. One said it need better time label.

- For 'Residents and Jobs' map, 8 of them showed interests on it, 1 of them showed interests on it and 1 people are not interested in it. People's purposes of using it are various. One wants to use it for traffic management on companies. And another uses it for making policy about bike routes.

As we can see, almost all visualizations are interested by different users but for different purposes. However, for some maps, it is not easy to understand what are expressed. This is the problems that some participants have met in task testing. These information could be provided to the development team of BikePRINT to consider which maps need better definitions and explanations. Also, some functionality were suggested by users that could be added in BikePRINT so that they will be more useful.

5.7.3. Interested analytical question

In this section, those questions that were chosen by more than half of users will be presented here.
1. What is the average number of cyclists on road segment X per day?
2. What is the average number of cyclists in city X per day?
3. Where do cyclist ride bikes in city X on day Y?
4. Which road segments are passed by cyclists per hour?
5. There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?
6. Where are the cycling trips whose origins are in city X and destinations are outside that city X?
7. How many people rode bikes from city X to other places outside that city on day Y?
8. What is the trend of movements between city X to city Y during the whole period?
9. Have the spatial distribution patterns in morning rush times of 5 working days changed in city X?
10. What is the similarity of spatial-temporal patterns of bike trajectories on the 5 working days in city X?
11. What is the difference of spatial-temporal patterns of movements in city X between working days and weekends?

On the other hand, there are some analytical questions that no one choose:
1. Where is the longest cycling trip in city X on day Y? Where are its origin and destination?
2. Where is the earliest/latest cycling trip in city X on day Y?
3. Where is the cycling trip that lasts for the longest time in city X on day Y?
4. What is the spatial-temporal distribution of cycling trajectories in the city which have the highest average speeds on day Y?

As we can see, these questions are more about information of one single trajectory so that it could be concluded that they more focus on the level of several trajectories.

5.7.4. Conclusion

By the questionnaire to all other users, only 11 users replied which is far less than expected numbers. It was informed that there are around 80 users were sent the questionnaire. Therefore, if more users can reply, more reliable results could be made.

5.8. Summary

In this chapter, the whole procedure of the user research was described in details including the set-up stage, implementation stage, results and the analysis of results. The purpose of this user research is to find limitations of BikePRINT and users' wishes of using it. In the user research, inviting some participants to complete tasks under the observation is the key part. From this testing, some recordings of their activities and voices are obtained. Then, I transcribed these recordings to the text document. Through applying the coding approach, I extracted the key information that reveal limitations from the documents. Users' performances are one of these important information. Then I compared and analyzed their performances and gave the summary of limitations. These will be considered with my previous findings to give improved visualization solutions for BikePRINT. In next chapter, I will describe my conceptual design of improvements according to the results of this chapter.

6. DESIGN OF IMPROVED VISUALIZATION SOLUTIONS AND RECOMMENDATIONS FOR BIKEPRINT

6.1. Introduction

This chapter will describe the design of the improved visualization solutions for BikePRINT according to the limitations found in the user research step and the self-evaluation. From the communication with Mr.Kruijf, who is one of development team of BikePRINT, I received an information that they are working on developing the new version of BikePRINT. Therefor, the improved solutions described in this chapter could be recommendations to them. In this chapter, first of all, the designed improved solutions will be proposed and explained in details. And then some participants' comments from the focus group that are related to these solutions will be mentioned as well.

6.2. Conceptual design

Through testing how participants interact with BikePRINT to complete tasks brought limitations that could be improved. The purpose of designing improved visualization solutions is to help them to explore cycling trajectory data, extract useful information and get insight in to them more efficiently. The improved solutions will focus on commonly used visualizations of them. Users should be able to use these visualizations to answer analytical questions with respect to spatial, temporal and attribute components of cycling trajectories and cycling infrastructures. In this section, I will first introduce designs of helping users to access/extract needed information. And then, the ways of visualizing them or improvements of existing visualizations of these information will be presented. As last, some functionality and features which could help users using BikePRINT with more good experience will be introduced. As users use the visualization 'intensity', 'absolute speed', 'delay' and 'routes' commonly, so the conceptual design focuses more on improving these visualizations.

6.2.1. Accessing needed data

In BikePRINT, users only can choose/select one road segment at a time by clicking it. If they want to analyze the information of multi-road segments, they could click them one by one and take notes but which is very inefficient and time-consuming. Or if they want to know the spatial patterns of road segments or trajectories with same value, they also cannot filter other irrelative data out. And users cannot extract information in a particular time interval in BikePRINT. So some questions with temporal information cannot be answered well by BikePRINT. So firstly I design the 'Value Filter' to help users choosing data with defined value/attributes, and then is the 'Time Filter' to help them select information in a time slot. In addition, another solution is called 'Spatial Searching' which helps users to select information within a particular area.

1) Value Filter (VF)

VF could be used to filter the road segments or delay points with the value that the user defined. VF boxes could be integrated in the interface of BikePRINT (Figure 6-1). In boxes, users can filter the segments they want. For example, if the user would like to show all road segments whose 'intensity' is more than 500, and 'speed' is less than 3 km/h, he/she can type these two numbers and all segments with

these two values will be shown. In the focus group discussion, participants proposed that sometimes they only want to show road segments with high/middle/low values on the map. Therefore, in this way, the urban planners can easily find where are road segments with high intensity and low average speed which could be roads that need improvements.
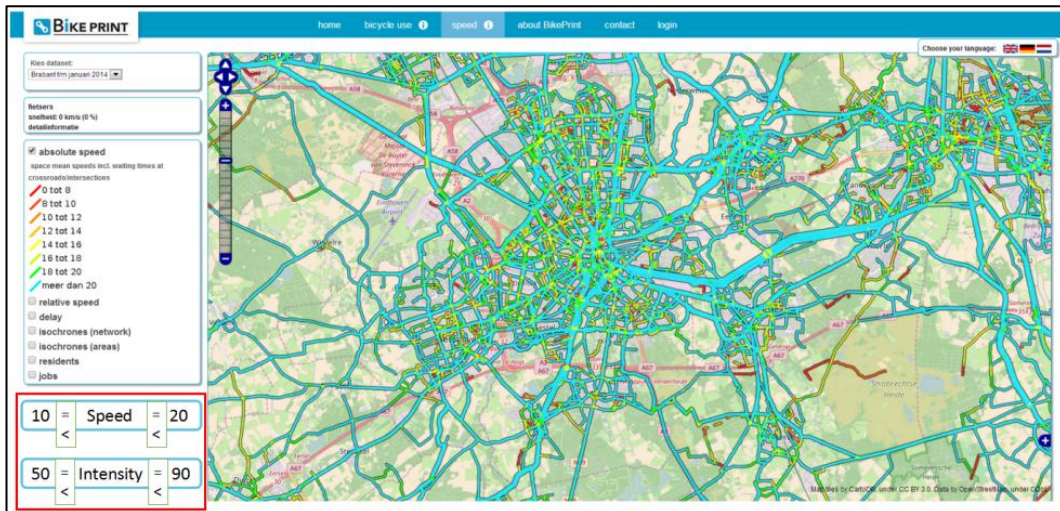
Figure 6-1 The interface that shows Value Filter of 'speed' map (in the red rectangle)

The VF also could be applied in 'Delay' points for users to find bottlenecks (Figure 6-2). So users can find all points with high waiting time and more delayed people, which is a definition of bottlenecks proposed by participants in the user research.
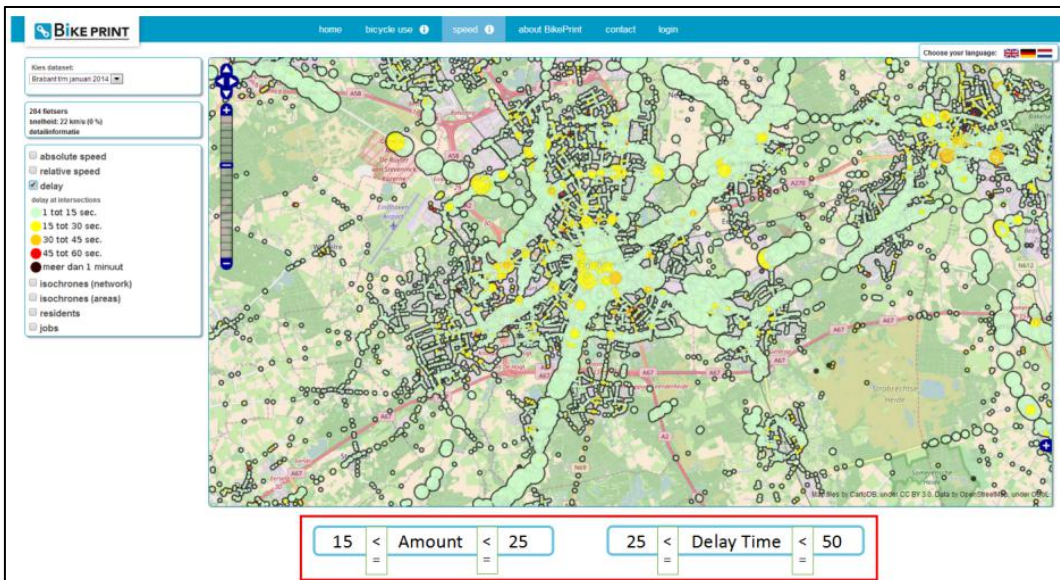
Figure 6-2 The interface that shows Value Filter of 'delay' map (in the red rectangle). In the box, 'amount' means the number of cyclists who delayed at this point and 'Delay Time' means the average delay time of this point.

2) Time Filter (TF)

TF could be applied to these four visualizations (intensity, speed, delay and routes) to select information within a time interval (e.g. all trajectories in morning rush hour). Two filter bars could be added in the interface, users can adjust time interval in it (Figure 6-3). The left one is used to filter data by time of the day, while the right one is used to filter data by days.
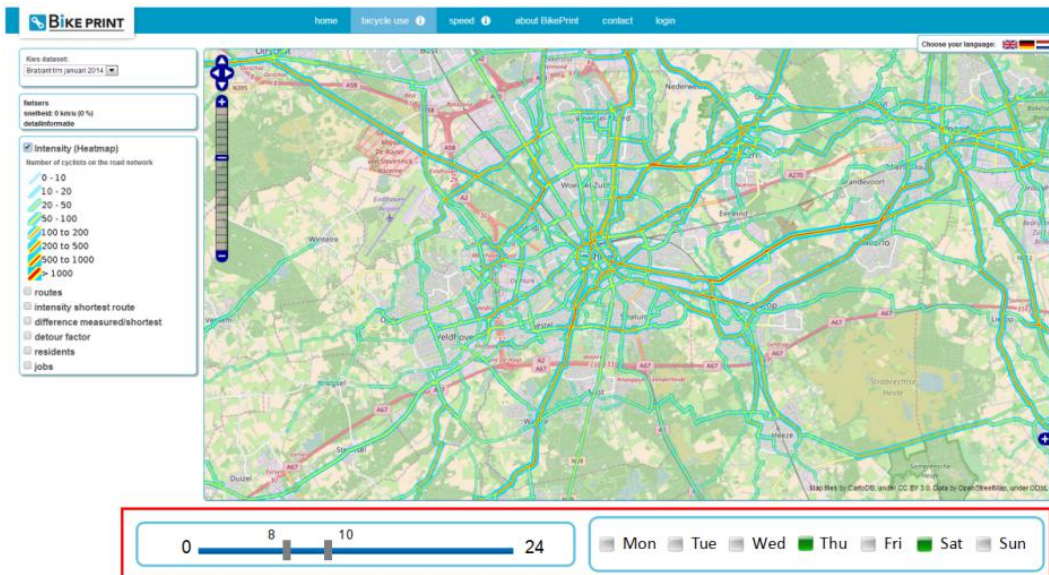
Figure 6-3 The interface that shows Time Filter of 'intensity' map (in the red rectangle). In the left, users can filter time of the day and in the right, they can choose days of the week.

3) Spatial Searching (SS)

SS also could be used in these four visualizations. However the ways of using it are different. Firstly, in 'intensity', 'speed' and 'delay' map, the road segments and points could be selected. So, users can use mouse to draw an area on the map where they are interested to know detailed information so that all segments/point within the area will be selected (Figure 6-4). Then, in the 'detailed information' box which is existing in BikePRINT, they can have these chosen elements' information.
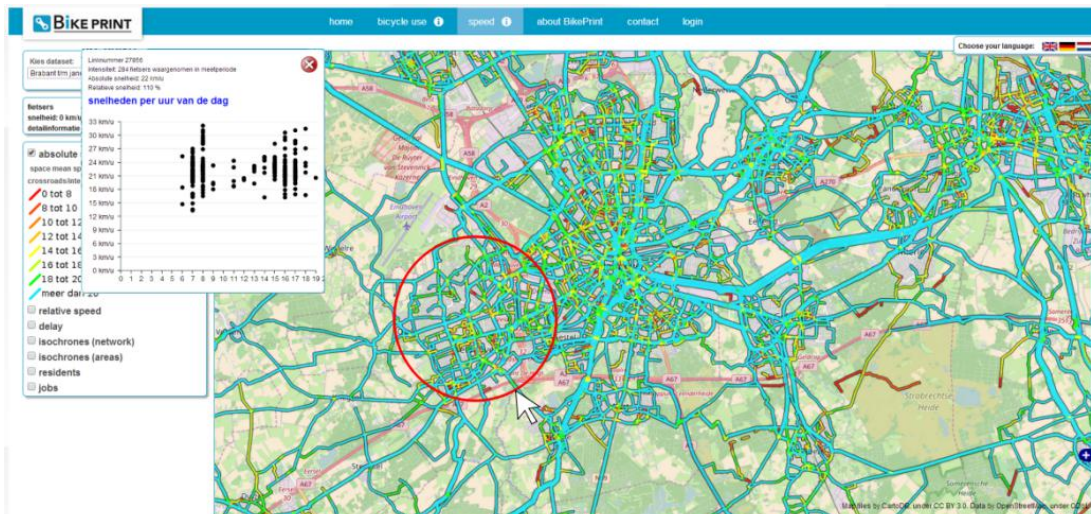
Figure 6-4 The interface that shows Spatial Searching of 'speed' map (in the red circle)

On the other hand, in BikePRINT, there is only one way to select cycling trajectories which is clicking a road segment. In this way, participants said that when the user click a road segment, actually she/he cannot exactly know which road segment is chosen because it won't be highlighted. So I designed an improvement with respect to this limitation. As shown by Figure 6-5(a), when the user moves the mouse on one road segment, it will be represent by yellow (or other colors). As shown by Figure 6-5(b), when the user clicks it, it changes to green color (or other colors). And then, all cycling trips that passed this segment will be shown and the user still knows which is the chosen segment (Figure 6-6).
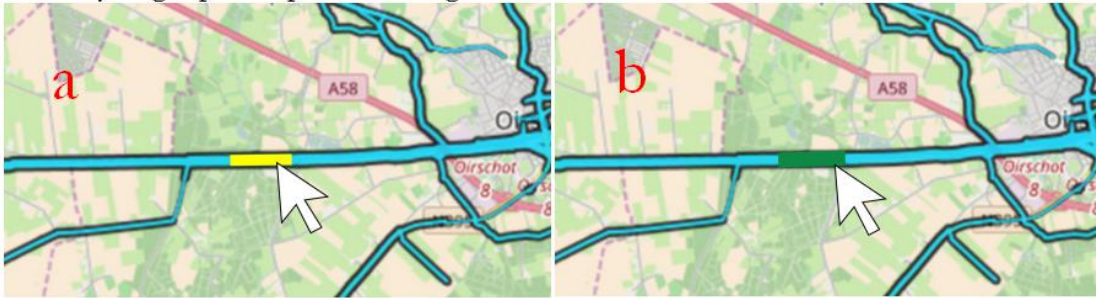
Figure 6-5 The functionality of using mouse to highlight and click the road segment
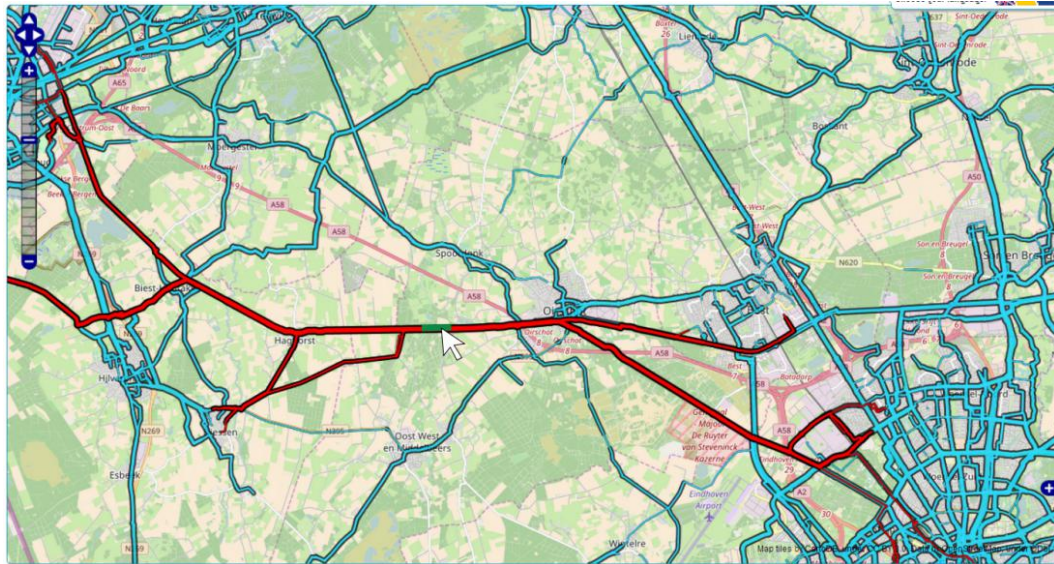
Figure 6-6 Cycling trips are shown but clicking the road segment. When the user clicks the road segment, it turns to green and all cycling trips that passed this segment are shown as well.

However, this way of selecting trajectories is obviously not enough. So I designed another way called 'selecting OD' that users can define two areas as the ranges of origins and destinations (Figure 6-7(a)). As shown in Figure 6-7(a), users can firstly click 'Origin' in the box in the bottom left hand corner and draw an area 'A' of origins. And then they can click 'Destination' in the same box to draw another area 'B' of destinations. As the result, all cycling trips whose origins are in area 'A' and destinations are in area 'B' will be selected and visualized on the map (Figure 6-7(b)). They also could only draw one area of origins or destinations. If the user click the button 'Both', all cycling trips whose origins and destinations are in area 'A' and 'B' will be selected and visualized on the map. Therefore, by this means, the analytical questions like 'what is the spatial patterns of cycling trajectories from city X to city Y?' could be answered. Actually, at first, I designed to clicking to road segment instead of drawing two areas to select cycling trips and then I presented it to participants in Amsterdam. They commented that using this way may leads to that there are no trajectories could be chosen. They preferred choosing two areas. Then, I discussed this with Mr.Kruijf in the meeting in Utrecht. He suggested that some interesting points could be set in BikePRINT and linked to its adjacent road segments. Therefore, users could choose interesting points as the origin or destination. In this way, for example, the questions 'Where are the destinations of cyclists from the central station of city X?' could be answered in which the central station is the interesting point. Therefore, drawing two areas and selecting interesting points are recommended to be designed.
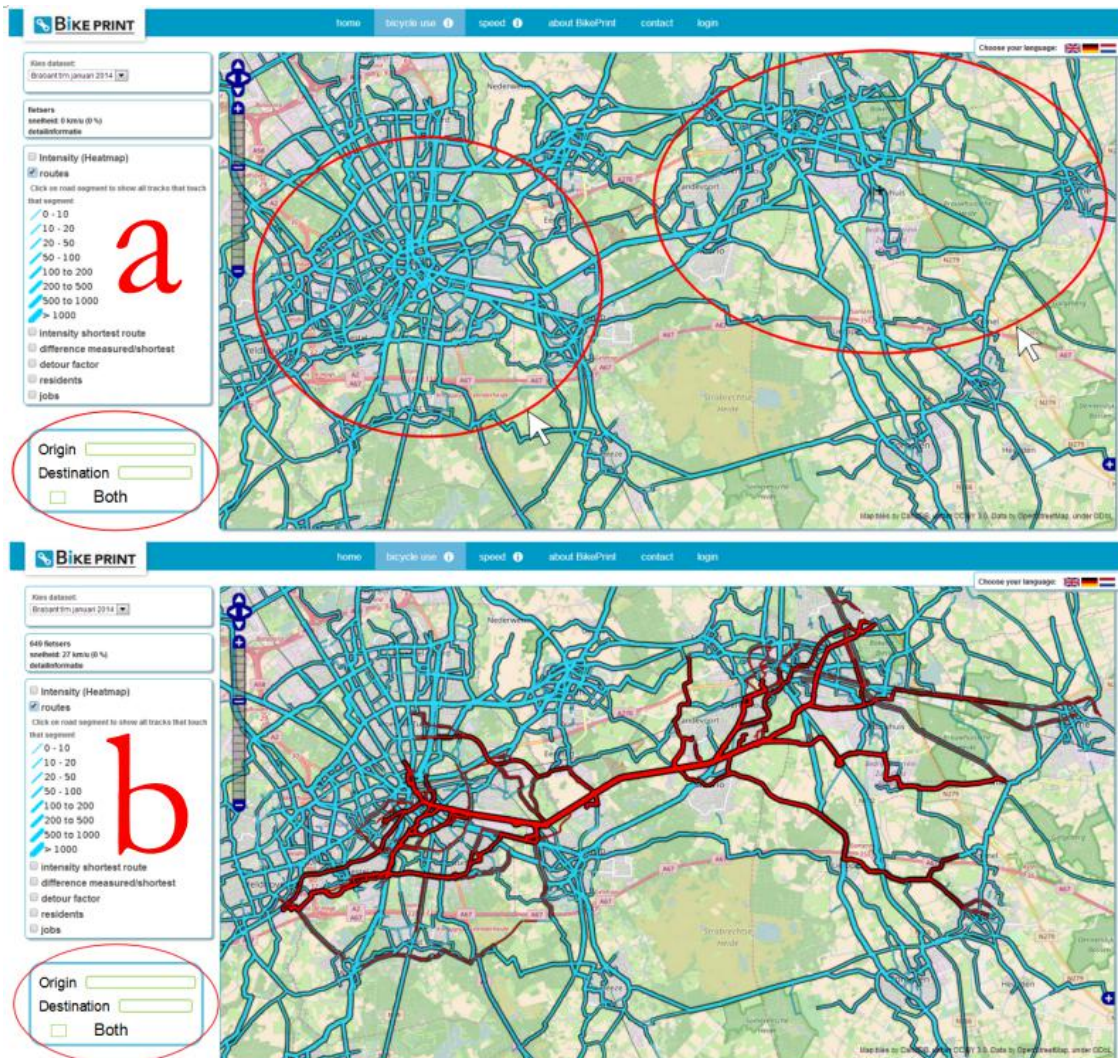
Figure 6-7 The example to show how to choose two area as origins and destinations and visualize chosen trajectories

6.2.2. Visualizations of accessed data

When the user accessed the data in BikePRINT, it is necessary to visualize these data in BikePRINT in suiTable ways. In, BikePRINT, there are two issues of visualizing data: (1) BikePRINT cannot provide visualization of some information (e.g. Origin, destination and direction of the trajectory); (2) some visualizations in BikePRINT are not desirable (e.g. The base map and sometimes users cannot distinguish the elements). At first, the improved solution is for the base map and the visual variables of the road segments.

1) The base map and Visual variables of road segments

In BikePRINT, the colorful base map effects users' reading of the information to a certain extent. In order to improve this, I looked through some other visualizations. In some other examples of visualization, produces used the base map with dark color and unmixed colors. Therefore, users can quickly focus on the layers with information. So I recommend that in new BikePRINT, the base map could be the subdued color and the colors of layers could be bright so that users will be less effected by the base map. However sometimes users may need the colorful base map when they want to take land use information into consideration, so a switch button is added and user can switch on/off the base map.

In addition, the different visual variables of road segments could be used to represent 'intensity' and 'speed' information together in which the size of line could be used to represent the amount of cyclists and the lightness of a color could be used to show the speed because some users wish to be able to analyze them. Figure 10 and Figure 11 are two examples for this design.

Figure 6-8 Map of London Cycle Hire. In the map, the thickness of road lines is used to show amount of journeys, while the lightness of red color is used to represent numbers of pollution index. (URL 10)

Figure 6-9 Visualization of cycling use in Philadelphia. In this map, the lightness of color and size visual variables of road lines are used to represent the amount of cycling (URL 11)

2)   Visualization of Trajectories

Cycling trips (trajectories) are one of most important information in BikePRINT which could be used to analyze cycling behaviors, identify spatial patterns, etc. However, in BikePRINT, some important information of trajectories cannot be visualized. For example, the origin/destination and direction of trajectories are not provided which were proposed by users in important to users. And when many cycling trips overlaps, users cannot distinguish them. They also cannot select one cycling trip from many trips to get the detailed information. Therefore, I designed some solutions for improving them as follows.

● **Multi-Trajectories**

This solution is designed to help users distinguish many overlapped trajectories on the map. Every single trajectory is represented by an unique color, while the overlapped part (overlapped segments) of them is shown by grey color. The intensity of grey color is used to represent how many trips overlap in this segment. This idea was inspired by the map of bus routes in Palma so I use this map to take the example (Figure 6-10). In the Figure, only one intensity of grey is used to show overlapped roads. As shown in Figure 6-11, users can also move the mouse on the road to know the numbers of bus. So in the design, users can move the mouse on one segment to know the amount of trajectories that passed.
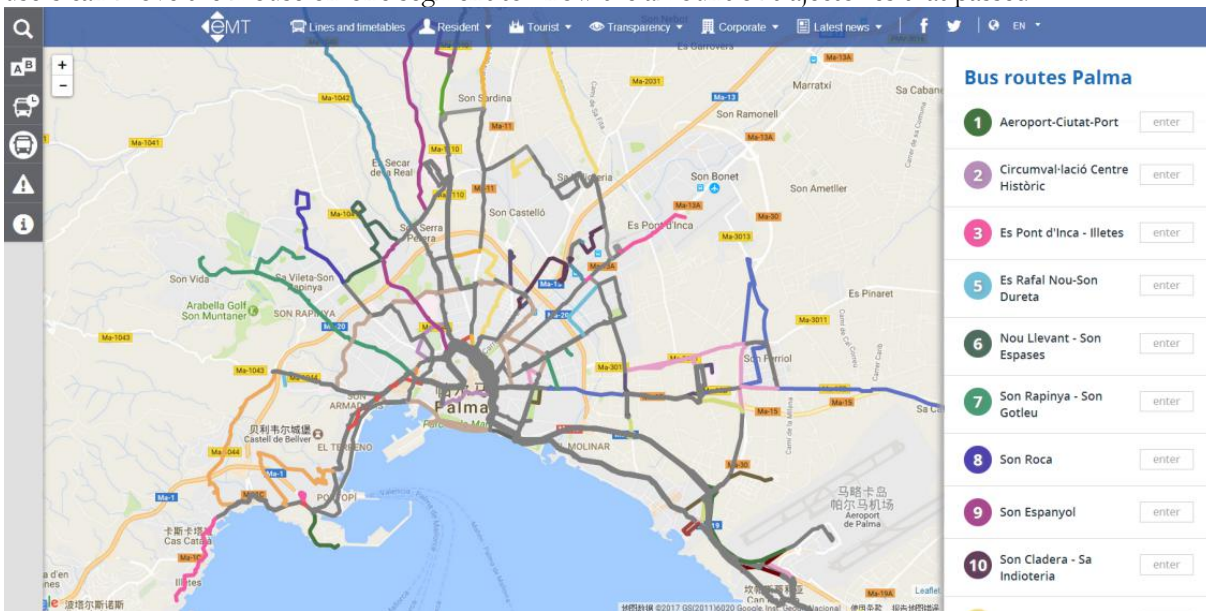
Figure 6-10 The example to show how to visualize multi-trajectories on the map (URL 12)
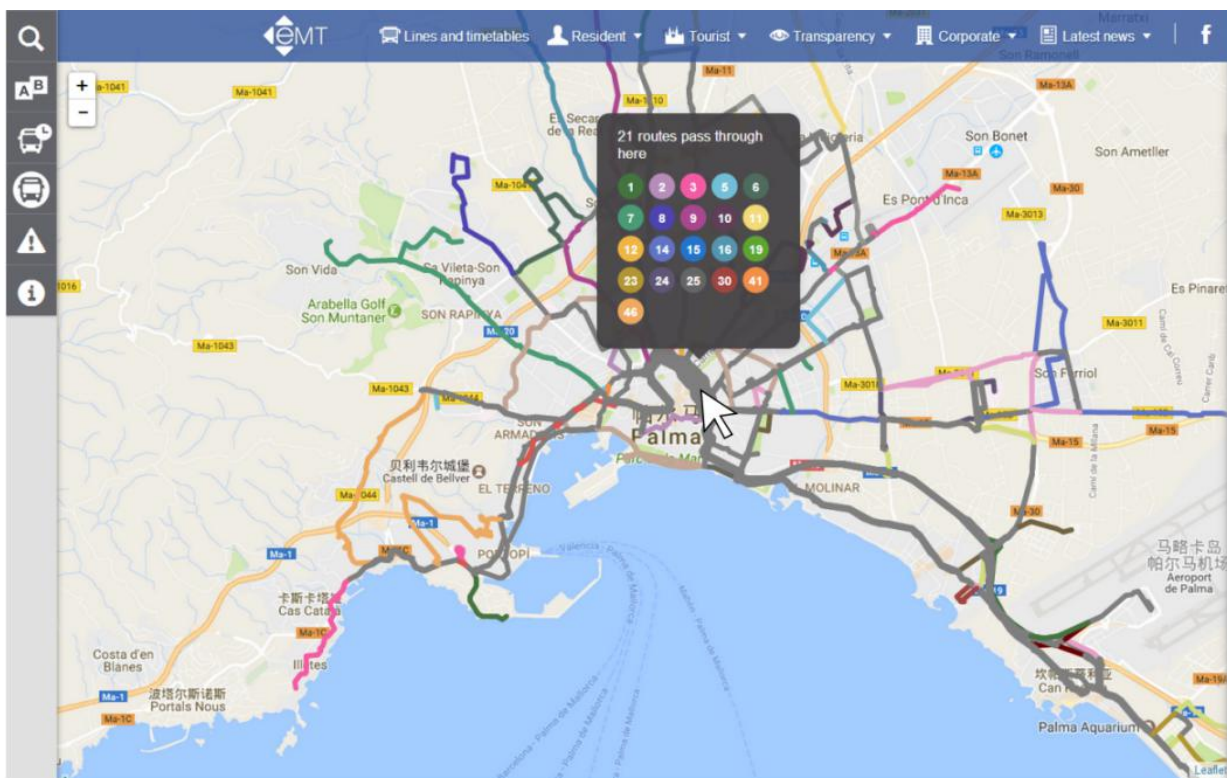
Figure 6-11 When the user click the road, all bus routes that pass this road will be shown (URL 12)

● **Directions of trajectories**

In BikePRINT, the direction of trajectories cannot be visualized. The participants proposed that they need these information to do the analysis in the focus group discussion. So I designed that arrows of lines are used in BikePRINT to show the direction. I drew some arrows on the map of BikePRINT to show this idea (Figure 6-12). In the Figure, a and b represent single direction of the trajectory but by different ways (a: inside the line; b: in the middle of the line). And c represents double direction of the trajectory.

Figure 6-12 Examples of using arrows to show directions

In this way of visualizing directions, the visual variable size could be applied to arrows to represent the amount of cyclists who used this road, as shown in Figure 6-13. Figure 6-14 is an example of the overview of using arrows to show directions on the map. There is another functionality that could be added in to help users know the amount of the arrow better. As we can see, when the user moves the mouse on an arrow, the numbers of passed cyclists is also shown.
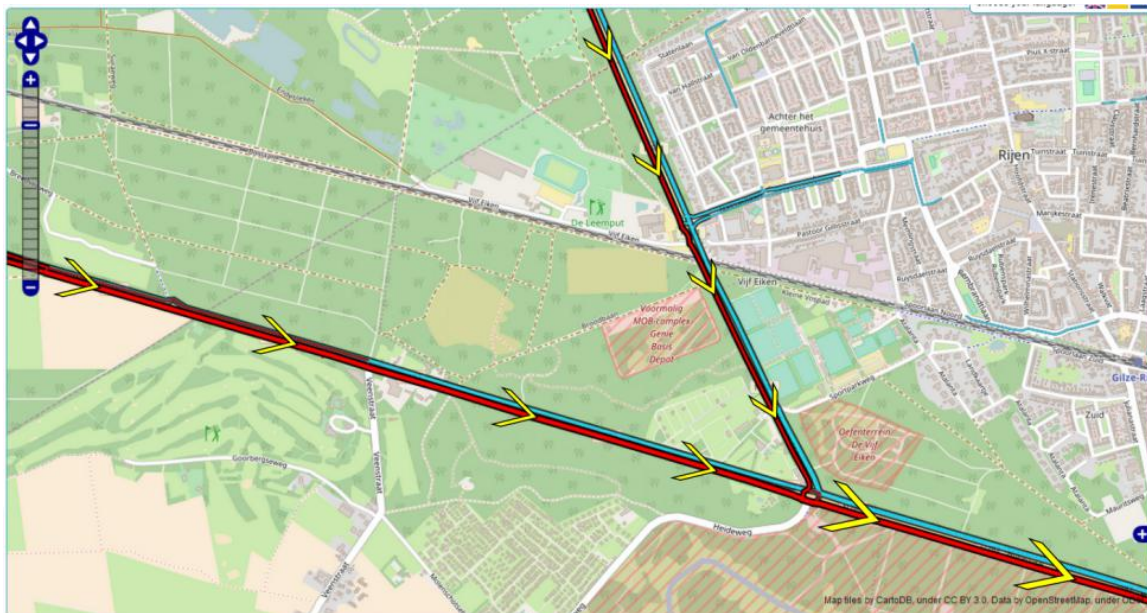
Figure 6-13 The size of arrows represent amounts
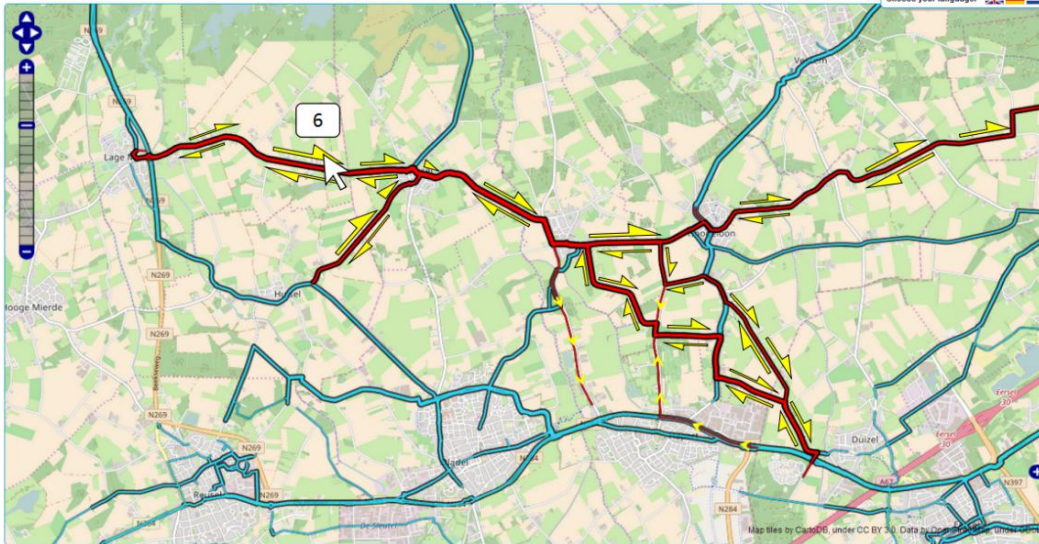
Figure 6-14 The overview of using arrows to show directions on the map

However, when I presented this solution to users, they considered that when the number of trajectories is large or when roads are close to each other (like roads in Amsterdam), this way may generates some clutters. And they proposed that for them, it is enough to only know the direction information of one road segment. However, I still keep these solutions because from the results of questionnaire from other users, some of them indicated that they need to analyze directions of routes or cycling trips. In order to meet Amsterdam peoples' needs, another solution was designed to meet this requirement. As we see in Figure 6-15, when the user move the mouse on one segment, the arrows with numbers are shown. So the user could know how many cyclists passed this segment from two directions. It could be more useful to combine with the Time Filter (TF).
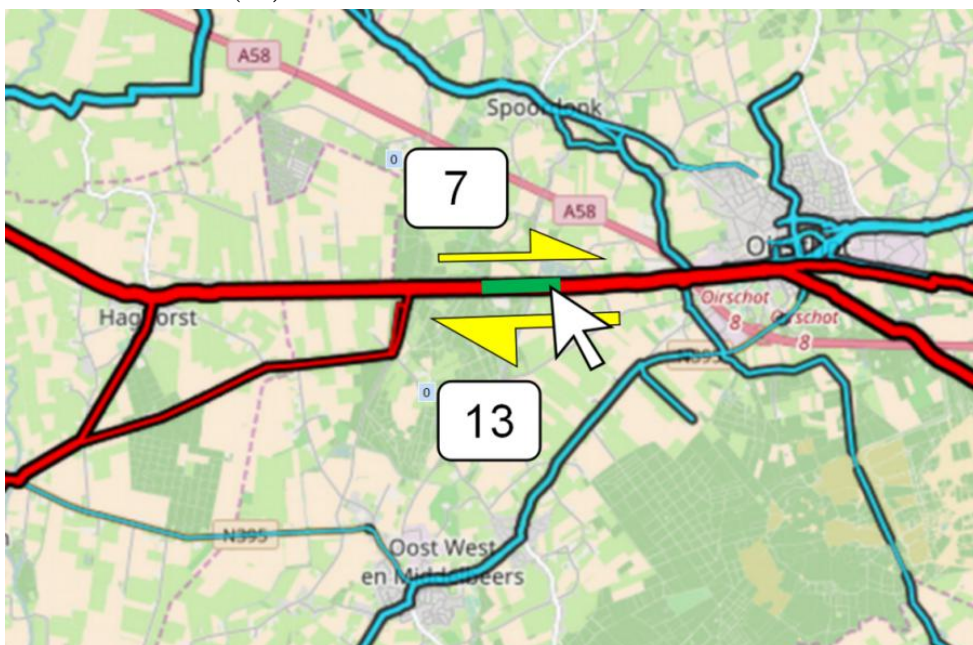
Figure 6-15 Users can move the mouse on the road segment to know the amounts of two directions

- **Visualization of Origins and Destinations**

In BikePRINT, when there are many trajectories are shown, some of them are overlapped. Users cannot make sure where are these trajectories start and end. However, information of origins and destinations of trajectories is useful for urban planners. For example, they could use these information to analyze that for

all trajectories which passed a particular road segment, where are their origins? Some cycling behaviors could be extracted from these information. Or for the analytical question 'Where are destinations of all trajectories whose origin is central station?'. Figure 6-16 is an example of visualizing the distribution of all cycling trajectories whose origin is in Eindhoven and destination is in Helmond area. Orange points are origins and black points are destinations, which could be more useful to combine with the Time Filter (TF).
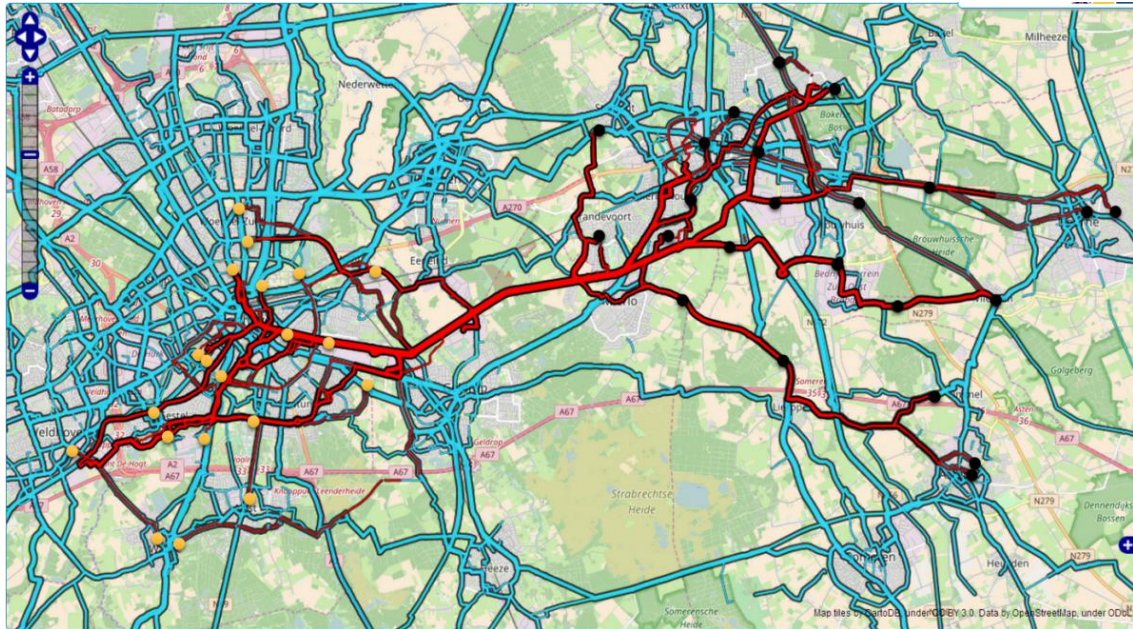
Figure 6-16 The example of visualizing the distribution of origins and destinations

- **Selecting one single trajectory to get detailed information**

As presented before, in BikePRINT, cycling trips can be represented by different colors and overlapped parts of them are represented by grey color. So the users can distinguish them in general. However, some of them may want to know detailed information of one of them. Therefore, I designed solutions for them to select one single trajectory from them. When the user clicks the unoverlapped part of one cycling trip, its whole trajectory will be highlighted and its detailed information will be shown in the 'detailed information' box like in Figure 6-2 of BikePRINT, as shown by Figure 6-17. In the Figure, yellow line is the chosen trajectory.In addition, users also are able to selecting one cycling trip by clicking its origin or destination (Figure 6-18).
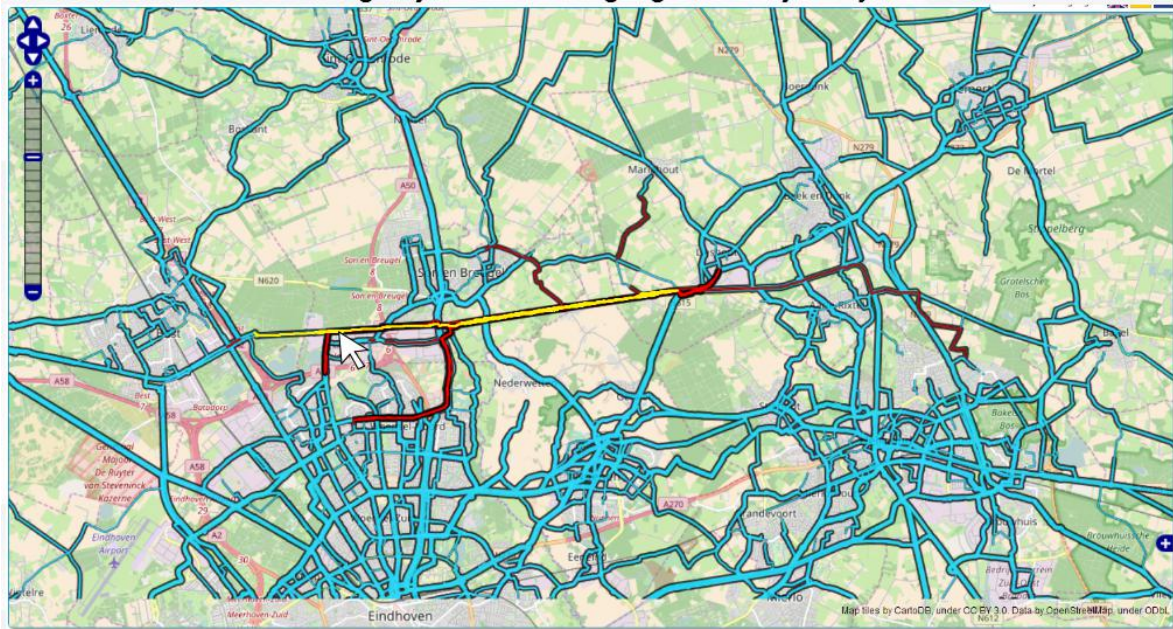
Figure 6-17 Users can move the mouse on the road segment to highlight it
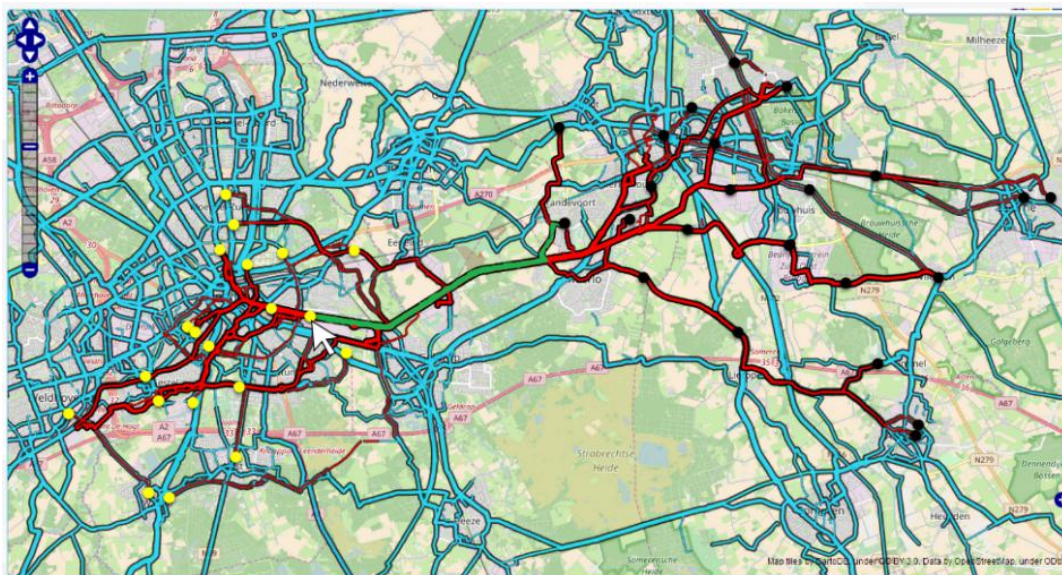
Figure 6-18 Users can move the mouse on the origin/destination to highlight the cycling trip

- **Visual variables of trajectory**

When the user select a trajectory and highlight it, it is shown by a solid color in which users cannot extract information. As introduced in chapter 4, in BikePRINT, the trajectory is represented by a sequence of road segments. Therefore, attributes of segments of this trajectory could be visualized as well by thickness of color visual variables of segments (line). As shown in Figure 6-19, for single trajectory, arrows are used to show its direction and thickness of color are used to show attributes of segments that this trajectory passed. This way could be used together with the previous solution.
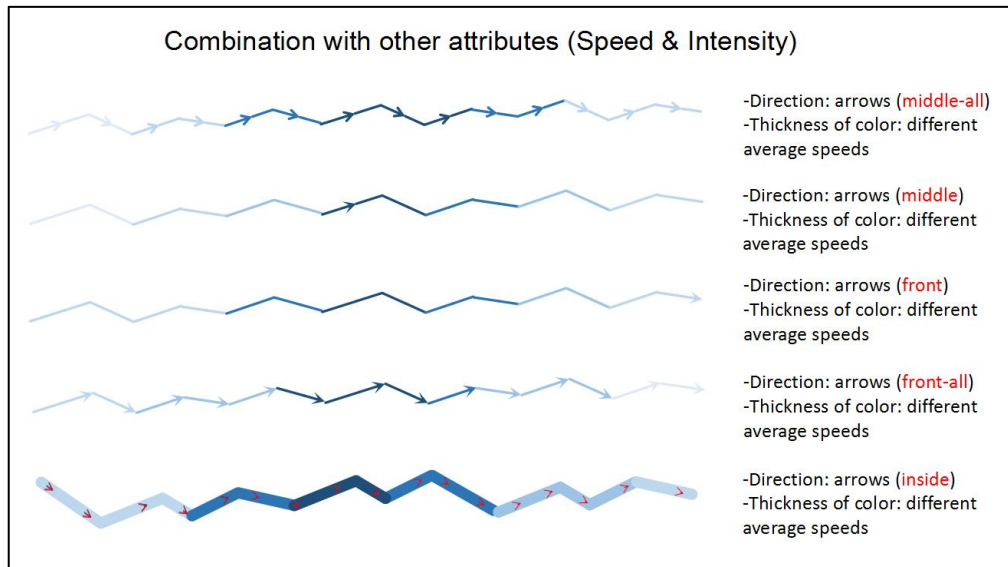
Figure 6-19 Examples of visualizing attributes in the trajectory

3) Small multiples

From users' interested questions and wishes in the section 5.4.1, I found that that they need to compare different information (e.g. Information in different time intervals). Therefore, the solution 'Small multiple' could be used to compare information with different spatial components, temporal components or attributes. By the help of this solution, some temporal and overall questions could be answered. For example, in Figure 6-20, user can compare different different patterns of cycling trips of 4 chosen road segments.
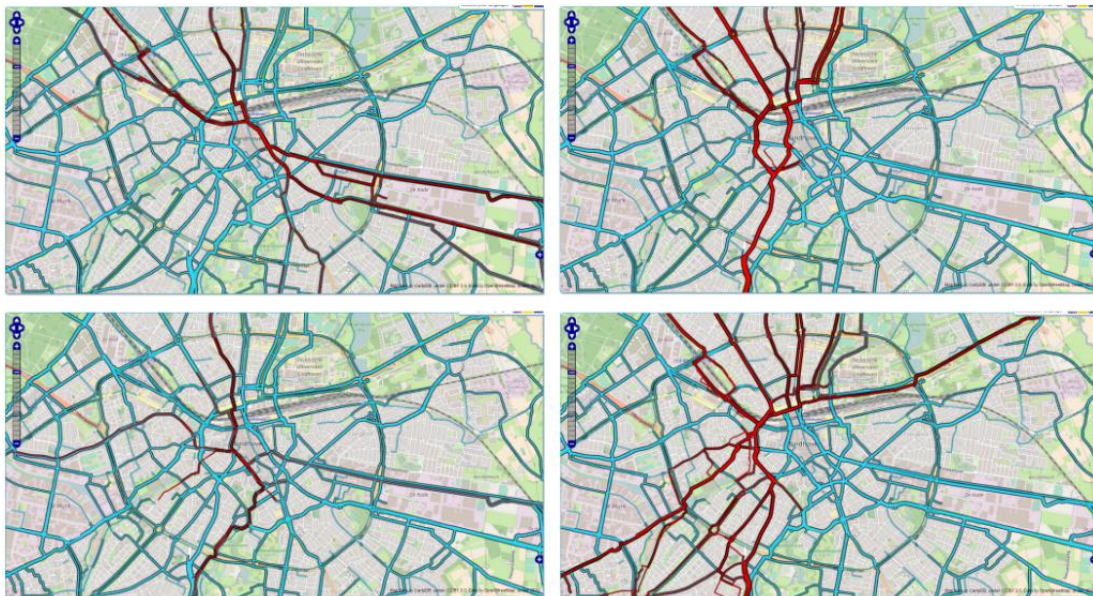
Figure 6-20 Example of small multiples to compare different patterns of cycling trips

6.2.3. Recommended functionality

In BikePRINT, some functionality could be added. These functionality help users using the application with better experience. Firstly, a search bar could be added into BikePRINT. Therefore, users can type names of cities, roads or Zip codes to locate them quickly. In addition, I recommend that all road segments which belong to a road could be linked to this road. Therefor, users could type the name of the road to select these road segments with convenience. Secondly, the calculator could be added in BikePRINT. Because urban planners are main users of BikePRINT, sometimes they needs the average

numbers or total numbers of elements' attributes (e.g. Average speed of road segments in an area) which was proposed by the feedback of questionnaires in the section 5.7.2. They can use the calculator to calculate these numbers and make the report. Thirdly, in BikePRINT, users' operations are irreversible which leads they cannot 'undo' some operations. Therefore, BikePRINT should be able to record their operations and provide this functionality. Fourthly, there should be a signal in BikePRINT to show its status so that users can know whether the BikePRINT is loading/processing or not. The last one is that it should be possible to open more than one window at a time for the small multiples.

6.2.4.   Summary

In this section, some possible visualization solutions are described from three aspects: accessing data, visualizing data and other functionality. By these solutions, users can extract spatio-temporal information of cycling trajectory data better than using current version of BikePRINT. And these data can be visualized by proposed solutions to meet their analytical purposes and assist them answer some analytical questions well. Therefore, these improved visualization solutions could be recommended to developers of BikePRINT to take into account when they are developing the new generation of BikePRINT.

6.3.   Discussion of solutions with users and the developer.

6.3.1.   Discussion with users

After I designed these solutions, I firstly presented these solutions to participants in Amsterdam when I conducted the focus group discussion with them. The PPT slides was used to demonstrate these designs (appendix). In the presentation, I made some animations to show how can they interact with new solutions to exploring data and help them understanding these solutions better. For each solution, participants provided their idea and comments. Some of their comments of each solution were described in the section 6.2. For most solutions, they thought these will be helpful to them exploring and analyzing cycling trajectory data in BikePRINT if these could be implemented. And then, I revised and optimized some solutions which they thought are not very helpful and added some new solutions. For example, they proposed that they do not need the solution of Figure 6-12 so I designed another solution shown by Figure 6-14 to meet their needs, which was described in the section 6.2.2. But I still keep the solution of Figure 6-12 because from the feedback of other users, they indicated that they may need this solution.

6.3.2.   Discussion with the developer

After that, on 16th March 2017, I met with Mr. Joost de Kruijf who is one of development team to discuss the improved solutions and share ideas. At first, I presented the summary of found limitations of BikePRINT and Mr.Kruijf agreed with these limitations. As for the functionality of selecting the road instead of the road segment, he indicated that it is a good way but for analytical part, purposes of analysis should be considered. Then, He said that Time filter is also could be used in 'Delay' map. Also for Time filter, he suggested that the days also could be filtered as well so that users can further compare the different information of weekdays and weekends. As to the solutions of using arrows to show directions, he though it may generate some clutters if there are many cycling trips. Another recommendation he provided is that the visual variables of road segments also could be used to show the percentage of amounts of cyclists. Another important thing is that multiple window are already possible in BikePRINT but only accessed with log-in account. It could be developed to be a comparison functionality. As the solution introduced in Spatial Searching, Mr.Kruijf suggested that interesting points could be used to select elements.

Beside to his comments and suggestions of my proposed solutions, he introduced some their new ideas of improving BikePRINT. Firstly, they designed a functionality to selecting cycling trips that users can draw a line on the map and all trajectories that pass this line will be shown. Secondly, they designed the additional view to show the detailed information of the road segment. When the user click the road segment, the 'detailed box' could show its detailed information including the direction. And they plan to design the dashboard where summary of detailed information could be reported. And he thought that dynamic visualization (e.g. Animations) could provide good visualization but difficult to interpret information and agreed with that urban planners need more practical and useful visualizations than fancy visualizations. According to these comments, my final possible solutions were proposed in this chapter.

6.4.   Conclusion

In this chapter, improved visualization solutions of BikePRINT were designed to assist users exploring and analyzing cycling trajectory data. By these solutions, users can access spatio-temporal information that they need to answer analytical questions. And then proposed visualization solutions help them to show accessed data so that they can intuitively analyze these information. Especially, the directions of cycling trips can be represented by arrows which show how cyclists moved on the road network. In addition, the distribution of origins and destinations for cycling trips could be visualized as well that helps users analyzing some cycling behaviors, such as finding where are common destinations of a particular interesting point as the origin (e.g. Central station). Beside to these two types of improved solutions, some functionality were recommended to developers which mainly means to improve users' experience of using BikePRINT. During the discussion with users and developers, they proposed many comments about these solutions. After that, final solutions were determined.

7. CONCLUSION

7.1. Introduction

Cycling trajectories data reveal cyclists behaviors and conditions of cycling road network usage. Through applying visualization methods, these data could be represented and meaningful information could be extracted and analyzed by analysts. However, unsuiTable and insufficient ways of visualizing these data cannot assist users to obtain an insight of the data. BikePRINT is one of this issue so it was focused as the case study in this research. Therefore, the purpose of this research is to design improved visualization solutions of these data through discovering and analyzing the limitations of BikePRINT. User-centered design plays the essential role in the research in which the usability testing of the application was implemented with actual users of BikePRINT or people in the related domain. In order to make the testing works well, some preparation works were made beforehand and the detailed plan of the testing was made. And the analytical tasks used in usability testing were formulated according to the feedback of pre-questionnaire. Then, the testing was implemented with two group of users. The whole testings was recorded by the software which generated audio/video and screen recordings. These recordings contain users activities and performances which were analyzed afterwards. Through the analysis, not only some limitations of BikePRINT were obtained, different performances of users were also known. In addition, users' needs and interested analytical questions were acquired and analyzed as well. In the mean time, ways of representing and visualizing trajectory data were reviewed by reading literature. Moreover, some ideas of improving visualizations of cycling trajectory data were inspired by some other examples of visualizing trajectory data from websites and literature. Then, the improved visualization solutions were made according according to results of these works from three aspects: accessing data, visualization and functionality. By the help of new solutions, users can interact with BikePRINT to extract needed information and visualizing them for their purposes. These solutions were presented to users and the developer of BikePRINT and they provided their preferences and comments. According to these comments solutions were revised and then described in the thesis. Theoretically, these solutions could assist users to answer analytical questions with respect to spatial, temporal information and attributes of cycling trajectories.

7.2. Revisiting the research questions

In section 1.2.2, research questions of this research were proposed. Answers to these questions will be summarized in this section.

Q1 What are spatio-temporal cycling trajectory data?
Spatio-temporal cycling trajectory data describe how cyclists moved in cities during the period of time and their path of movements, which is one type of movement data.

Q2 What are the characteristics of these data?
Characteristics of these data could be described from three aspects: spatial, temporal and attributes characteristics. For spatial characteristics, cyclists' positions (including origins and destinations), speeds, directions, traveled distance and others could be described. For temporal characteristics, duration of

movements, starting/ending time and others are provided. Attribute characteristics of these data could be properties of cyclists, happened events in the movement and environments.

Q3 What are the design solutions of visual representation of these data?
From the section 2.4, as the way of recording the trajectory, it could be visually represented by a sequence of points and lines on the map. The visual variables of these points and lines are used to represent attributes of the trajectory. In addition, in some case, the sequence of road segments where the trajectory passed in the road network also are used to represent the trajectory.

Q4 What technical tools have been used in visualizing these data?
The answer to this question could be found in the section 2.5. There are four types of tools that could be used to visualize these data: (1) the static map, (2) small-multiple maps and (3) animated maps and (4) the space-time cube. In addition, Visual Analytics method that combines some visualizations and interactive methods helps people to explore spatio-temporal cycling trajectory data.

Q5 What are the characteristics and (dis)advantages of these solutions?
1) Using static maps is one of the earliest ways of visualizing trajectory data in which the visual variables are used to represent changes in different times, represent an event that occurs at a certain time moment or show the properties of moving objects in the map.however, it is less useful for representing temporal components. And if the dataset is big, some visual clutters may generate so that the functionality filtering or spatial generalization could be used to prevent this.
2) Small-multiple maps is a way of combining several static maps that respectively represent different information in different certain moments. Every small map could represent a moving state of objects in a trajectory in different times. However, if the amount of moments/small maps increases, the size of each small map deceases that will make people looking at the information in maps with more difficulties size that some detailed information may lose.
3) The animation map represents the change of the movement by displaying a sequence of images one by one in a single view. There are two main types of animated maps: time-based or non-time-based. However, using the animated map is less effectively and time consuming if there are big amount of moments needed to be dealt with.
4) Space-Time Cube (STC) is a developed visualization method that uses the three-dimensional visualization to represent both spatial and temporal information of movements within a cube, where the x and y axis represent the geographic locations and the third dimension z-axis is used to represent time. It could help people to represent the space, time and their relationships of the movement at the same time. However, because people's perceptual issues towards to 3D properties, when dealing with massive trajectories, some clutters may be generated that leads to the undesirable visualization result. Therefore, interactive methods are always combined with the space-time cube for users to filter the data or manipulate the map, change the point of view of observing the visualization

In addition, in this research, the visualization solution BikePRINT is focused. BikePRINT provides several visualizations of trajectory data that could be used for different purposes. Although, after the usability testing of BikePRINT, some limitations are found as follows:
1) BikePRINT have poor visual representations of cycling trajectories (e.g. Their origins/destinations, directions, etc.).
2) BikePRINT visualizes data separately, not integrated so that users cannot analyze them together.
3) BikePRINT doesn't visualize temporal information well (e.g. Users cannot extract information of a particular time interval.).

4) BikePRINT provides less interactive methods to help users exploring the data (e.g. Users cannot filter data by some values or times.).

Q6 Who are the users of spatio-temporal cycling trajectory data and for which purposes do they use the data?

There are many types of users who can use these data for different purposes. For example, cyclists could use the data to analyze their behaviors. However, in this research, users are mainly urban transportation planners/analysts. As described in the section 5.4.1, they usespatio-temporal cycling trajectory data for analyzing information of cycling trajectories (e.g. speeds, directions, etc), cycling use in the infrastructures, amounts of cyclists and their behaviors.

Q7 What analytical questions do they want to find answers to ?

In section 4.4.3, I proposed some analytical questions of cycling trajectory data and send it to users to evaluate and choose. After the analysis of feedback from them, those questions are their interests:

- What is the average number of cyclists on road segment X per day?
- What is the average number of cyclists in city X per day?
- Where do cyclist ride bikes in city X on day Y?
- Which road segments are passed by cyclists per hour?
- There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?
- Where are the destinations of cyclists from the central station of city X during the morning rush time of day Y?
- What are the spatial-temporal patterns of trajectories from city X to city Y in the morning rush time and night rush time of working days? What is the difference?
- Where are the cycling trips whose origins are in city X and destinations are outside that city X?
- How many people rode bikes from city X to other places outside that city on day Y?
- What is the trend of movements between city X to city Y during the whole period?
- Have the spatial distribution patterns in morning rush times of 5 working days changed in city X?
- What is the similarity of spatial-temporal patterns of bike trajectories on the 5 working days in city X?
- What is the difference of spatial-temporal patterns of movements in city X between working days and weekends?

On the other hand, there are some analytical questions that they are not interested:

- Where is the longest cycling trip in city X on day Y? Where are its origin and destination?
- Where is the earliest/latest cycling trip in city X on day Y?
- Where is the cycling trip that lasts for the longest time in city X on day Y?
- What is the spatial-temporal distribution of cycling trajectories in the city which have the highest average speeds on day Y?
- On which day did most cyclists depart from the central station in city X?
- What are the spatial-temporal patterns of cycling trajectories from the east to west part in city X on day Y?
- What are the spatial-temporal patterns of cycling trajectories in the city X during the whole period?

Q8 Which analytical questions cannot be answered well with the existing visualization solutions?

In the section 4.6.3, I evaluated all proposed questions whether could be answered with the help of BikePRINT. Of these interested questions in the answer to Q7, only the fifth question (There is a place

outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?) could be answered. Others cannot be answered well.

Q9 What are the problems they meet when using the existing visualization solutions for their use cases?
This question could be answered by contents of the section 5.6.3.
Problems could be summarized as follows:
1) Users cannot select more than one elements in BikePRINT to get detail information at a time.
2) Some users cannot understand some information of BikePRINT.
3) Users only can guess some information in BikePRINT but cannot know the exact amount.
4) Users cannot find some information in BikePRINT (e.g. Origins/destinations and directions of BikePRINT)

Q10 What factors lead to these problems?
- For the first: BikePRINT doesn't provide the functionality of selecting more elements.
- For the second: In some maps of BikePRINT, it doesn't give clear definition and explanation of what are expressed by it.
- For the third: BikePRINT doesn't provide functionality to filter data by the particular values or times. Therefore, users only can guess the answers.
- For the forth: BikePRINT doesn't provide visualizations of these information so users cannot find them.

Q11 According to the user requirements and existing solutions' limitations, what could be the design of improved solutions for the visualization of the data?
The answers to this question could be found in Chapter 6. I designed solutions from three aspects: Accessing needed data, Visualizations of accessed data and Other functionality.

- For Accessing needed data: Value Filter, Time Filter and Spatial Searching are three main ways of accessing the data. Value Filter helps users filtering data by amounts of attributes, Time Filter helps users filtering data by times and Spatial Searching helps users filtering data by the particular areas.

- For Visualizations of accessed data: solutions of the base map and visual variables of road segments are proposed as first. Then visualizations of trajectories are designed to show multi-trajectories, directions, origins and destinations and the way of selecting one single trajectories to get detail information. Then, the visualization of combining trajectories and attributes of road segments were designed. In the end, Small multiples was proposed for users to compare different information.

- And other functionality were recommended to improve users' experience of using BikePRINT. Firstly, a search bar could be added into BikePRINT. Secondly, the calculator could be integrated in BikePRINT. Thirdly, the recommendation is that BikePRINT should be able to record their operations. Fourthly, there should be a signal in BikePRINT to show its status so that users can know whether the BikePRINT is loading/processing or not. The last one is that it should be possible to open more than one window at a time for the small multiples.

Q12 What visualization tools can be used to design the visualization solution?

As the design was made on the basis of existing visualizations in BikePRINT, accordingly to the answers to Q11, static maps with interactive methods and functionality are designed so that users can manipulate the map to extract and visualization information. Then, small multiples was recommended in solutions so that user can use it to compare different information. As the design was not really implemented and recommended to the develop team of BikePRINT, it should be practical for them to develop.

7.3.   Some limitations of the research

Firstly, because of the time issue, only five users have participated in usability testing of the user research step, although they have brought many findings of BikePRINT. It is better if I can invited more people to do the tasks by interacting with BikePRINT, which may generate more comprehensive results. In addition, I only received 11 replies of the questionnaire. If there are more, the results could be more reliable. Then, designed solutions actually are not implemented to prototypes because of some reasons. And because I have no opportunity to have a meeting with technical people of development team of BikePRINT, it is unknown whether these solutions are practical for them to develop or not. Therefore, usability of these new solutions cannot be known.

7.4.   Recommendations for future work

This research mainly focus on 'Fietstelweek' data, application 'BikePRINT' and users of them. Therefore, some further works can focus on other visualization environments of cycling trajectories data if there are. When there are more users, more wishes of using the application can be acquired so that solutions could be designed more suiTable. Moreover, it is better if these solutions could be discussed with more users and developers. And if these solutions could be implemented, more further works can be made.

LIST OF REFERENCES

Adrienko, N., & Adrienko, G. (2011). Spatial generalization and aggregation of massive movement data. *IEEE Transactions on visualization and computer graphics, 17*(2), 205-219.

Ahmed, M., & Wenk, C. (2012). *Constructing street networks from GPS trajectories.* Paper presented at the European Symposium on Algorithms.

Aigner, W., Miksch, S., Müller, W., Schumann, H., & Tominski, C. (2007). Visualizing time-oriented data—a systematic view. *Computers & Graphics, 31*(3), 401-409.

Amar, R., Eagan, J., & Stasko, J. (2005). *Low-level components of analytic activity in information visualization.* Paper presented at the Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on.

Andrienko, G., & Andrienko, N. (2009). Visual analytics for geographic analysis, exemplified by different types of movement data *Information Fusion and Geographic Information Systems* (pp. 3-17): Springer.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., Kisilevich, S., & Wrobel, S. (2011). A conceptual framework and taxonomy of techniques for analyzing movement. *Journal of Visual Languages & Computing, 22*(3), 213-232.

Andrienko, G., Andrienko, N., Bak, P., Keim, D., & Wrobel, S. (2013). *Visual analytics of movement*: Springer Science & Business Media.

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., . . . Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science, 24*(10), 1577-1600.

Andrienko, G., Andrienko, N., Kopanakis, I., Ligtenberg, A., & Wrobel, S. (2008). Visual analytics methods for movement data *Mobility, data mining and privacy* (pp. 375-410): Springer.

Andrienko, G., Andrienko, N., Schumann, H., & Tominski, C. (2014). Visualization of trajectory attributes in space–time cube and trajectory wall *Cartography from Pole to Pole* (pp. 157-163): Springer.

Andrienko, G. L., & Andrienko, N. V. (1999). Interactive maps for visual data exploration. *International Journal of Geographical Information Science, 13*(4), 355-374.

Andrienko, N., & Andrienko, G. (2006). *Exploratory analysis of spatial and temporal data: a systematic approach*: Springer Science & Business Media.

Andrienko, N., & Andrienko, G. (2007). Designing visual analytics methods for massive collections of movement data. *Cartographica: The International Journal for Geographic Information and Geovisualization, 42*(2), 117-138.

Andrienko, N., & Andrienko, G. (2013). Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization, 12*(1), 3-24.

Andrienko, N., Andrienko, G., & Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing, 14*(6), 503-541.

Andrienko, N., Andrienko, G., Pelekis, N., & Spaccapietra, S. (2008). Basic concepts of movement data *Mobility, data mining and privacy* (pp. 15-38): Springer.

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J., & Kretchmer, U. (2002). Testing the usability of interactive maps in CommonGIS. *Cartography and Geographic Information Science, 29*(4), 325-342.

Barbosa, L., Kormaksson, M., Vieira, M. R., Tavares, R. L., & Zadrozny, B. (2014). *Vistradas: Visual Analytics for Urban Trajectory Data.* Paper presented at the GeoInfo.

Beecham, R. (2015). Using bikeshare datasets to improve urban cycling experience and research urban cycling behaviour. *Cycling Futures: From Research into Practice*, 267.

Beecham, R., & Wood, J. (2014a). Characterising group-cycling journeys using interactive graphics. *Transportation Research Part C: Emerging Technologies, 47*, 194-206.

Beecham, R., & Wood, J. (2014b). Exploring gendered cycling behaviours within a large-scale behavioural data-set. *Transportation Planning and Technology, 37*(1), 83-97.

Beecham, R., Wood, J., & Bowerman, A. (2012). *A visual analytics approach to understanding cycling behaviour.* Paper presented at the Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on.

Beecham, R., Wood, J., & Bowerman, A. (2014). Studying commuting behaviours using collaborative visual analytics. *Computers, Environment and Urban Systems, 47*, 5-15.

Bertin, J. (1983). Semiology of graphics: diagrams, networks, maps.

Bevan, N., Carter, J., & Harker, S. (2015). *ISO 9241-11 revised: What have we learnt about usability since 1998?* Paper presented at the International Conference on Human-Computer Interaction.

Biadgilgn, D., Blok, C., & Huisman, O. (2011). Assessing the cartographic visualization of moving objects. *Momona Ethiopian Journal of Science, 3*(1), 80-104.

Borgnat, P., Fleury, E., Robardet, C., & Scherrer, A. (2009). *Spatial analysis of dynamic movements of Vélo'v, Lyon's shared bicycle program.* Paper presented at the ECCS'09.

Boyandin, I. (2013). *Visualization of temporal origin-destination data.*

Boyandin, I., Bertini, E., Bak, P., & Lalanne, D. (2011). *Flowstrates: An Approach for Visual Exploration of Temporal Origin‐Destination Data.* Paper presented at the Computer Graphics Forum.

Buja, A., Cook, D., & Swayne, D. F. (1996). Interactive high-dimensional data visualization. *Journal of computational and graphical statistics, 5*(1), 78-99.

Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE computer graphics and applications, 25*(4), 12-16.

Chen, Z., Shen, H. T., & Zhou, X. (2011). *Discovering popular routes from trajectories.* Paper presented at the Data Engineering (ICDE), 2011 IEEE 27th International Conference on.

Chuah, M. C., & Roth, S. F. (1996). *On the semantics of interactive visualizations.* Paper presented at the Information Visualization'96, Proceedings IEEE Symposium on.

Come, E., Randriamanamihaga, A., & Oukhellou, L. (2014). *Spatio-temporal usage pattern analysis of the Paris Shared Bicycle Scheme: a data mining approach.* Paper presented at the Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment.

Come, E., Randriamanamihaga, N. A., Oukhellou, L., & Aknin, P. (2014). *Spatio-temporal Analysis of Dynamic Origin-Destination Data Using Latent Dirichlet Allocation: Application to Vélib'Bike Sharing System of Paris.* Paper presented at the TRB 93rd Annual meeting.

Corry, M. D., Frick, T. W., & Hansen, L. (1997). User-centered design and usability testing of a web site: An illustrative case study. *Educational Technology Research and Development, 45*(4), 65-76.

Dix, A. (2009). *Human-computer interaction*: Springer.

Dodge, S., Weibel, R., & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information Visualization, 7*(3-4), 240-252.

Donald, N. (1988). The design of everyday things. *Doubled Currency.*

Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*: Intellect books.

Elmqvist, N., & Fekete, J.-D. (2010). Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE Transactions on visualization and computer graphics, 16*(3), 439-454.

Fisher, P., Dykes, J., & Wood, J. (1993). Map design and visualization. *The Cartographic Journal, 30*(2), 136-142.

Frank, A. U. (1998). Different types of "times" in GIS. *Spatial and temporal reasoning in geographic information systems*, 40-62.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). *Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?* Paper presented at the Proceedings of the SIGCHI conference on Human Factors in Computing Systems.

Georgy, J., & Noureldin, A. (2011). Tightly coupled low cost 3D RISS/GPS integration using a mixture particle filter for vehicular navigation. *Sensors, 11*(4), 4244-4276.

Giannotti, F., & Pedreschi, D. (2007). Geography, mobility, and privacy: a knowledge discovery vision: Springer, Heidelberg.

Giannotti, F., & Pedreschi, D. (2008). Mobility, data mining and privacy: A vision of convergence *Mobility, data mining and privacy* (pp. 1-11): Springer.

Goesele, M., Grosch, T., Preim, B., Theisel, H., & Toennies, K. (2012). The Great Wall of Space-Time.

Gonçalves, T., Afonso, A. P., & Martins, B. (2014a). Visualization Techniques of Trajectory Data: Challenges and Limitations.

Gonçalves, T., Afonso, A. P., & Martins, B. (2014b). *Visualizing human trajectories: comparing space-time cubes and static maps.* Paper presented at the Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014-Sand, Sea and Sky-Holiday HCI.

Gonçalves, T., Afonso, A. P., & Martins, B. (2015). Cartographic visualization of human trajectory data: overview and analysis. *Journal of Location Based Services, 9*(2), 138-166.

Gonçalves, T., Afonso, A. P., Martins, B., & Gonçalves, D. (2013). *ST-TrajVis: interacting with trajectory data.* Paper presented at the Proceedings of the 27th International BCS Human Computer Interaction Conference.

Greaves, S., & Figliozzi, M. (2008). Collecting commercial vehicle tour data with passive global positioning system technology: Issues and potential applications. *Transportation Research Record: Journal of the Transportation Research Board*(2049), 158-166.

Guo, D. (2009). Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Transactions on visualization and computer graphics, 15*(6).

Guo, D., & Zhu, X. (2014). Origin-destination flow data smoothing and mapping. *IEEE Transactions on visualization and computer graphics, 20*(12), 2043-2052.

Guo, H., Wang, Z., Yu, B., Zhao, H., & Yuan, X. (2011). *Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection.* Paper presented at the Pacific Visualization Symposium (PacificVis), 2011 IEEE.

Huang, X., Zhao, Y., Ma, C., Yang, J., Ye, X., & Zhang, C. (2016). TrajGraph: A graph-based visual analytics approach to studying urban network centralities using taxi trajectory data. *IEEE Transactions on visualization and computer graphics, 22*(1), 160-169.

Jiménez, P., Nogal, M., Caulfield, B., & Pilla, F. (2016). Perceptually important points of mobility patterns to characterise bike sharing systems: The Dublin case. *Journal of Transport Geography, 54*, 228-239.

Kjellin, A., Pettersson, L. W., Seipel, S., & Lind, M. (2010a). Different levels of 3D: An evaluation of visualized discrete spatiotemporal data in space-time cubes. *Information Visualization, 9*(2), 152-164.

Kjellin, A., Pettersson, L. W., Seipel, S., & Lind, M. (2010b). Evaluating 2d and 3d visualizations of spatiotemporal information. *ACM Transactions on Applied Perception (TAP), 7*(3), 19.

Koua, E. L., MacEachren, A., & Kraak, M. J. (2006). Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science, 20*(4), 425-448.

Kraak, M.-J. (2003a). Geovisualization illustrated. *ISPRS journal of photogrammetry and remote sensing, 57*(5), 390-399.

Kraak, M.-J. (2003b). *The space-time cube revisited from a geovisualization perspective.* Paper presented at the Proc. 21st International Cartographic Conference.

Kraak, M.-J., & Madzudzo, P. (2007). *Space time visualization for epidemiological research.* Paper presented at the ICC 2007: Proceedings of the 23nd international cartographic conference ICC: Cartography for everyone and for you.

Kraak, M.-J., & Ormeling, F. (2011). *Cartography: visualization of spatial data*: Guilford Press.

Kraak, M. J. (2008). Geovisualization and time–new opportunities for the space–time cube. *Geographic visualization: concepts, tools and applications*, 293-306.

Kveladze, I. (2015). *Space - time cube design and usability.*

Kveladze, I., Kraak, M.-J., & Van Elzakker, C. P. (2015). The space-time cube as part of a GeoVisual analytics environment to support the understanding of movement data. *International Journal of Geographical Information Science, 29*(11), 2001-2016.

Kwan, M.-P., & Lee, J. (2004). Geovisualization of human activity patterns using 3D GIS: a time-geographic approach. *Spatially integrated social science, 27*.

Macedo, J., Vangenot, C., Othman, W., Pelekis, N., Frentzos, E., Kuijpers, B., . . . Theodoridis, Y. (2008). Trajectory data models *Mobility, Data Mining and Privacy* (pp. 123-150): Springer.

Miller, H. J. (2005). A measurement theory for time geography. *Geographical Analysis, 37*(1), 17-45.

Ostermann, F. O. (2010). Digital representation of park use and visual analysis of visitor activities. *Computers, Environment and Urban Systems, 34*(6), 452-464.

Peuquet, D. J. (1994). It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of american Geographers, 84*(3), 441-461.

Phan, D., Xiao, L., Yeh, R., & Hanrahan, P. (2005). *Flow map layout.* Paper presented at the Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on.

Qian, L., Wachowicz, M., Peuquet, D., & MacEachren, A. (1997). *Delineating operations for visualization and analysis of space-time data in GIS.* Paper presented at the Gis/lis.

Ren, L., Du, Y., Ma, S., Zhang, X., & Dai, G. (2014). Visual Analytics Towards Big Data. *Journal of Software, 25*(9), 1909-1936. doi:10.13328/j.cnki.jos.004645

Robinson, A. C., Chen, J., Lengerich, E. J., Meyer, H. G., & MacEachren, A. M. (2005). Combining usability techniques to design geovisualization tools for epidemiology. *Cartography and Geographic Information Science, 32*(4), 243-255.

Roth, R. E. (2012). Cartographic interaction primitives: Framework and synthesis. *The Cartographic Journal, 49*(4), 376-395.

Scheepens, R., Willems, N., van de Wetering, H., & Van Wijk, J. J. (2011). *Interactive visualization of multivariate trajectory data with density maps.* Paper presented at the Pacific Visualization Symposium (PacificVis), 2011 IEEE.

Shneiderman, B. (1996). *The eyes have it: A task by data type taxonomy for information visualizations.* Paper presented at the Visual Languages, 1996. Proceedings., IEEE Symposium on.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data & knowledge engineering, 65*(1), 126-146.

Spencer, R. (2007). Information Visualization: Design for Interaction. *Person Education.*

Tobler, W. R. (1987). Experiments in migration mapping by computer. *The American Cartographer, 14*(2), 155-163.

Tominski, C., Schumann, H., Andrienko, G., & Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. *IEEE Transactions on visualization and computer graphics, 18*(12), 2565-2574.

Tweedie, L. (1997). *Characterizing interactive externalizations.* Paper presented at the Proceedings of the ACM SIGCHI Conference on Human factors in computing systems.

van Elzakker, C. P. (2004). *The use of maps in the exploration of geographic data*: Utrecht University.

Vogel, M., Hamon, R., Lozenguez, G., Merchez, L., Abry, P., Barnier, J., . . . Robardet, C. (2014). From bicycle sharing system movements to users: a typology of Vélo'v cyclists in Lyon based on large-scale behavioural dataset. *Journal of Transport Geography, 41*, 280-291.

Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia-Social and Behavioral Sciences, 20*, 514-523.

Wikipedia. (2016). Retrieved from https://en.wikipedia.org/wiki/Geovisualization

Won, J.-I., Kim, S.-W., Baek, J.-H., & Lee, J. (2009). *Trajectory clustering in road network environment.* Paper presented at the Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on.

Wood, J., Slingsby, A., & Dykes, J. (2011). Visualizing the dynamics of London's bicycle-hire scheme. *Cartographica: The International Journal for Geographic Information and Geovisualization, 46*(4), 239-251.

Yi, J. S., ah Kang, Y., & Stasko, J. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on visualization and computer graphics, 13*(6), 1224-1231.

Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. *PloS one, 10*(10), e0137922.

URL 1. http://www.bikeprint.nl/fietstelweek
URL 2. http://www.bikeprint.nl/index.php
URL 3.
http://www.ee.ucr.edu/~mourikis/index.html?http%3A//www.ee.ucr.edu/%7Emourikis/project_pages /vision_INS.htm
URL 4. https://www.techsmith.com/camtasia.html
URL 5. http://atlasti.com/
URL 6. http://www.bikeprint.nl/fietstelweek/
URL 7. https://www.openstreetmap.org/#map=5/51.500/-0.100
URL 8. https://carto.com/
URL 9. http://otranscribe.com/
URL 10. http://mappinglondon.co.uk/2012/london-cycle-hire-and-pollution/
URL 11. http://www.dvrpc.org/webmaps/cyclephilly/
URL 12. http://www.emtpalma.es/en/lines-timeTables

APPENDIX 1

The questionnaire

Dear Sir/Madam,

We were informed by Mr. Joost de Kruijf (Utrecht University) that you are willing to participate in my research on the visualization of 'Fiets Telweek' data. Thank you very much for that.

My name is Shengce Wang. I am a Geoinformatics Master student in the Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente. My research is about designing visualization solutions for cycling trajectory data. I am using the 'FietsTelWeek' data as a case study and their current visualization through the BikePRINT application. The objective of my research is to recommend improvements of the visualization of 'FietsTelWeek' data through BikePRINT so that users get better answers to the questions they have regarding cycling in their study area.

I hope you are willing to be involved in my research through three successive steps:

1. I would first like to ask you to complete the questionnaire below and return it to me on or before 27 January 2017.

2. One week later I would like to come to Amsterdam and interview you individually. In the same session of about 30 minutes I would also like to observe you when you are interacting with BikePRINT. This observation will provide me with information on how the software is actually being used.

3. At the end of the day I would like to organize a Focus Group session with you and your colleagues to wrap up your needs and wishes with respect to the visualization of 'FietsTelWeek' data.

I hope you agree that I will make video and audio recordings of sessions 2 and 3. For academic reasons I will have to carefully analyze these recordings. Please be assured that in the resulting scientific reporting you will remain anonymous.

Finally, I regret to inform you that I do not master the Dutch language. I hope you do not mind to communicate with me in English and that you also answer the questionnaire below in the English language.

Once again: thank you very much for your time and cooperation!

QUESTIONNAIRE

1. Please provide your personal details:
-Name ………………………………………………...………………………………………
-Email ………………………………………………...…………………………………………

2. What is your educational background?
            ………………………………………………………………………………………………..

3. Please briefly describe your work and function (in particular in relation with cycling):
………………………………………………………………..…………………………………………
………………………………...……………………………………………………...…………………………
………………………………………..…………………………………………………………………………
…………………………………………..……………….……...…

4. Do you (plan to) work with 'FietsTelWeek' data? If so, for what purpose?
....
....

5. Do you (also) work with other cycling data? If so, which data and for what purpose?
…………………………………………………………………………….
…………………………………………………………………………….
…………………………………………………………………………….
…………………………………………………………………………….

6. Please indicate whether you work with cycling data for:
        -The municipality of Amsterdam only
        -Amsterdam and its surroundings
        -Amsterdam and one or more other municipalities
        -The Netherlands as a whole (including Amsterdam)
        -Or……………………………………………………

7. How often and for what purpose do you make use of maps (paper or digital) in your work?
        -Daily – Purpose: …………………………………………………………..
        -Several times per week – Purpose: ……………………………………………
        -Several times per month – Purpose: ……………………………………………
        -Several times per year – Purpose: ……………………………………………
        -I never make use of maps ……………………………………………………

8. Do you know the web application 'BikePRINT'?
    Yes
    No

9. If you know BikePrint, how often do /did you use it?
        -Daily
        -Several times per week
        -Several times per month
        -Only once or twice
        -I have never used it

10. If you have never used BikePrint, or only used it once or twice, what is the reason for that?
……
……………………………….……………………………………………………………………………………
……
……

11. There are many possible visualizations in 'BikePRINT'. Please tick those visualization
    options shown below that you have used before or you may want to use in future (You are
    interesting in it). If you have any remarks about them (e.g. For what purpose did/do you use

it? Did it solve your analytical questions and how? Or did it provide the information you are looking for and how many? And their limitations, your wishes and needs about them, etc.), please add them. If the visualization that you are not interested in, please skip it.

1) Intensity: This visualization shows the number of cyclists on the road network. Different colors mean different amounts of cyclists on the road segment.

☐ I have used it before.

Remarks: ....

....

☐ I may want to use it.

Remarks: ....

....

2) Routes: This visualization shows trajectories of cycling trips, the user can choose one road segment and click it to show all cycling trajectories that touch that segment. Blue lines are all cycling routes and red lines are shown routes of one chosen segment.

☐ I have used it.

Remarks: ....

....

☐ I may want to use it.

Remarks: ....

....

3) **Intensity shortest route** (fictive number of cyclists if everyone takes the shortest route according to the cycling route planner.) When the quality of a certain part of the cycle network is very low, most cyclists will probably avoid these routes. In a simple demand approach, these routes will not be improved on short notice because of the low number of cyclists. BikePRINT identifies these routes by comparing the actual observed cyclists on each segment with the fictive number of cyclists we would have if everyone would cycle the shortest route.

This visualization uses different colors to show the different fictive number of intensity.

☐ I have used it.

Remarks: ....

....
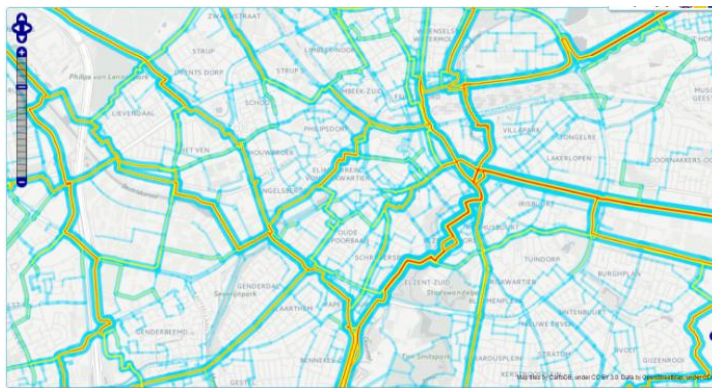
☐ I may want to use it.

Remarks: ....

....

4) **difference measured/shortest**: This visualization shows difference between actually measured number (actual number of cyclists) and fictive shortest routes before. Blue lines and red lines are opposite. Network segments where more actual cyclists are observed than would be expected from the shortest routes are coloured blue and interpreted as routes that are attractive enough that cyclists are willing to take a detour to use them. Conversely, the red coloured segment can be interpreted as "less attractive": cyclists appear to take detours in order to avoid these routes.

☐   I have used it.

Remarks: ....

....

☐   I may want to use it.

Remarks: ....

....

5) Detour factor (shows cycling detours compared to a straight-line connection)
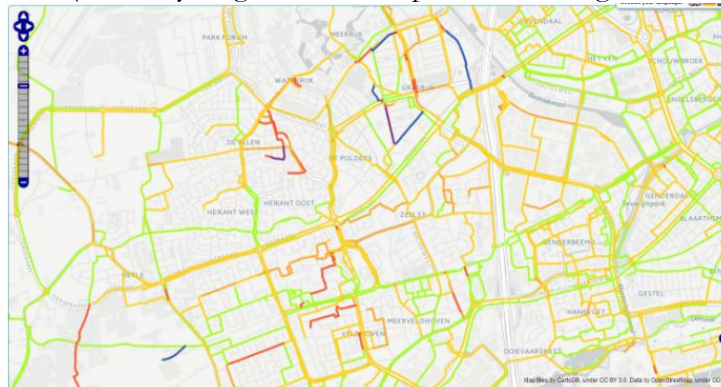
☐   I have used it.

Remarks: ....

☐   I may want to use it.

Remarks: ....

....

6) Absolute speed: shows average cycling speeds of each road segment including waiting
times at crossroads/intersections. Different colors are used to distinguish different
amounts.

☐ I have used it.

Remarks: ....

....
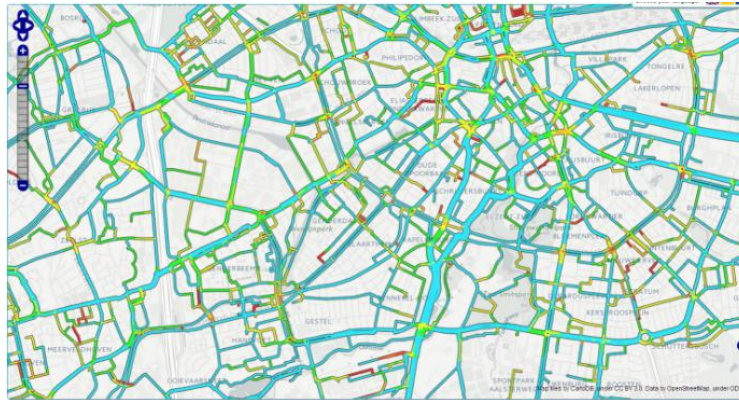
☐ I may want to use it.

Remarks: ....

....

7) **Relative speed:** This visualization shows the relative speed. The relative speed is the number of speed of the road segment that compared to the mean speed of all routes.

☐ I have used it.

Remarks: ....

....
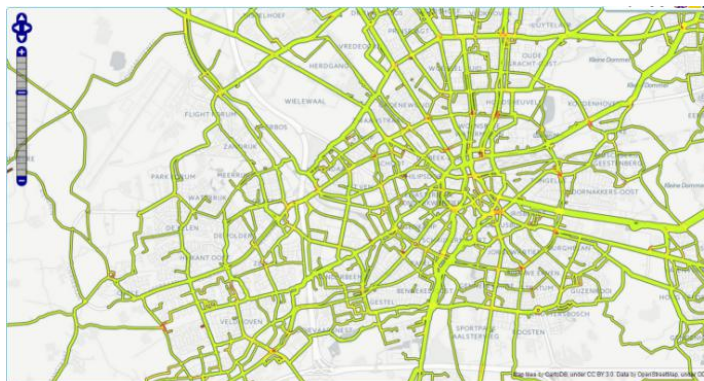
☐ I may want to use it.

Remarks: ....

8) **Delay time** (at intersections) the color of points represents the amount of delay time. The size of points represents the amount of people who delayed in this point.

☐ I have used it.

Remarks: ....

....

☐ I may want to use it.

Remarks: ....

....

9) Isochrones / network (click on a road segment to see travel times from that road
segment)

☐ I have used it.

Remarks: ....

....
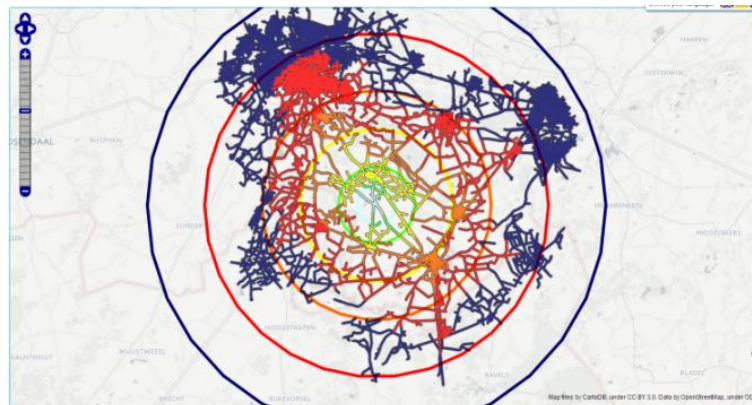
☐ I may want to use it.

Remarks: ....

....

10) Isochrones/areas (click on an area to see travel times from that area)

☐ I have used it.

Remarks: ....

....
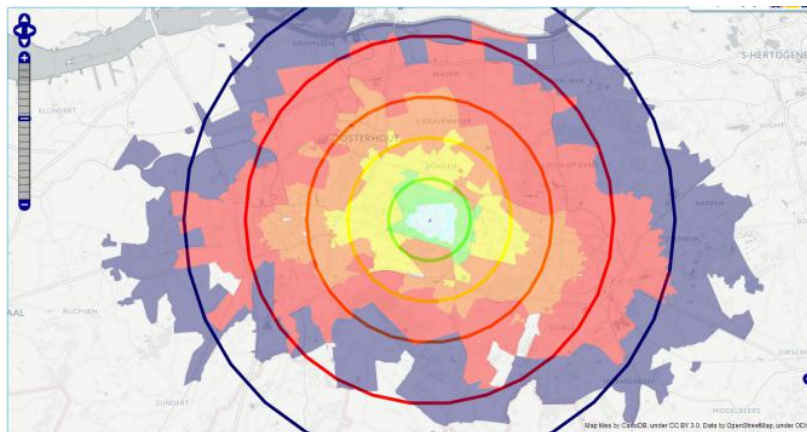
☐ I may want to use it.

Remarks: ....

....

Residents and Jobs: shows resident and job points on the map.

☐ I have used it.

Remarks: ....

....
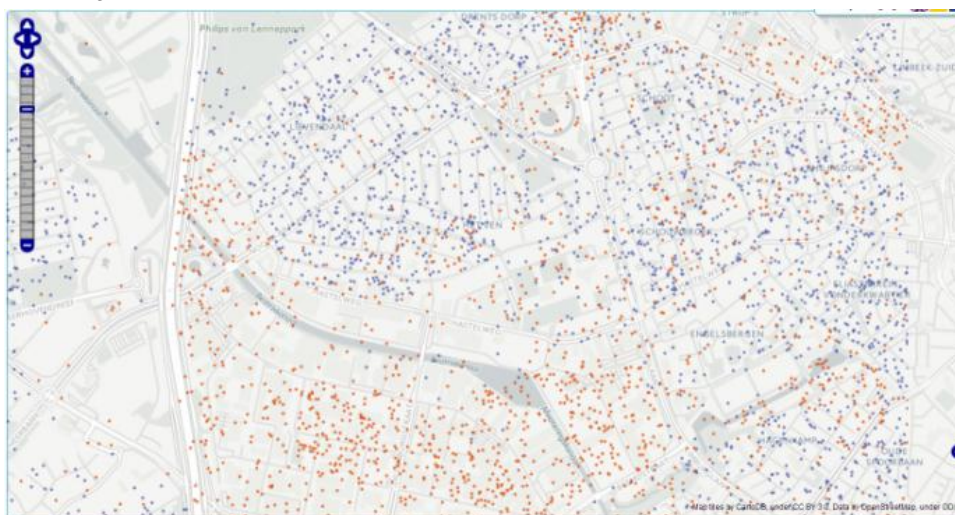
☐ I may want to use it.

Remarks: ....

....

12. Below you can find a list of spatio-temporal questions we think you would like to have answers to when analyzing cycling behaviour. Please tick the questions that you consider to be relevant.

Elementary questions:

1.What is the average number of cyclists on road segment X per day?

2.What is the average number of cyclists in Amsterdam per day?

3.Where do cyclist ride bikes in Amsterdam on day Y?

4.Which road segments are passed by cyclists per hour?

5.Which train stations did most cyclists depart from on day Y?

Intermediate questions:

1.Where is the longest cycling trip in Amsterdam on day Y? Where are its origin and destination?

2.Where is the earliest/latest cycling trip in Amsterdam on day Y?

3.Where is the cycling trip that lasts for the longest time in Amsterdam on day Y?

4.There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. Which route is most popular between the central station and this place?

5.Where are the cycling trips whose origins are in Amsterdam and destinations are outside that Amsterdam?

6.Where are the destinations of cyclists from the central station of Amsterdam during the morning rush time of day Y?

7.How many people rode bikes from Amsterdam to other places outside that city on day Y?

8.Where is the intersection that cyclists need to wait for the longest time?

9.Which road segments are chosen by cyclists most often on different days in Amsterdam?

Temporal questions:

1.What is the trend of movements from central station of Amsterdam to different destinations from morning to night on day Y?

2.What is the trend of movements between Amsterdam to city Y during the whole period?

3.There is a place outside of the city center where many students (Campus) or employees (the companies area) commute every day. What is the trend of cycling movements between the central station and this place?

4.What is the change of movements on 7 different days in Amsterdam?

5.On which day did most cyclists depart from the central station in Amsterdam?

6.Have the spatial distribution patterns in morning rush times of 5 working days changed in Amsterdam?

Overall questions:

1.What are the spatial-temporal patterns of cycling trajectories from the east to west part in Amsterdam on day Y?

2.What is the spatial-temporal distribution of cycling trajectories in the city which have the highest average speeds on day Y?

3.What are the spatial-temporal patterns of trajectories from Amsterdam to city Y in the morning rush time and night rush time of working days? What is the difference?

4.What is the similarity of spatial-temporal patterns of bike trajectories on the 5 working days in Amsterdam?

5.What is the difference of spatial-temporal patterns of movements in Amsterdam between working days and weekends?

6.What is the difference of spatial-temporal patterns of movements in Amsterdam between the morning rush time and night rush time on day Y ?

7.What are the spatial-temporal patterns of cycling trajectories in the Amsterdam during the whole period?

13. If you think there are any relevant questions missing (e.g. Your interested questions or questions that you had tried to solve by the help of 'BikePRINT' but failed), please list them below:

…………………………………………………………………………..…………………………………
………
…………………………………………………………………………..…………………………………

Please submit the completed questionnaire by email to s.wang-1@student.utwente.nl on or before 27 January 2017.

APPENDIX 2

Instruction and Task form

Dear participant,

- This user research aims to find out the limitations of the web application 'BikePRINT' and collect your suggestions and requirements for improving it. The emphasis is on the (cartographic) visualization of the Fietstelweek data. The experiment is observing you to do some analytical tasks that will be recorded. You will work in the environment of the web application 'BikePRINT'. The web application 'BikePRINT' provides several visualizations of cycling data by different web pages. In the meantime, you will get **7** analytical tasks to do, as indicated below. Please use 'BikePRINT' to try finding answers for these tasks. The test will take about 40 minutes. During the procedure, your voice, your facial expressions and your mouse activities will be recorded.

-If you meet some problems about the technical aspect or if you are confused about tasks before and during the testing, please do not hesitate to tell me. I will try my best to help you. However, if you meet some problems when dealing with the tasks (e.g. You cannot find the visualization page you are looking for), I cannot provide help because it is the content of the testing. If you feel tired please let me know, so we can take a break, and you can continue the tasks from the point where you stopped (the recording will then pause). Thanks for your understanding.

-During the procedure, **we expect you to think aloud (speak loudly)** whatever you are seeing, feeling, looking for and thinking about.

- When you finish all tasks, please say 'I finished all tasks'. So I will turn down all recorders and collect the papers.

- Now, if you have some unclearness about the procedure or 'BikePRINT', please ask me. If you don't have, please start with the tasks below.

Analytical tasks

These are some analytical tasks for you to complete by interacting with 'BikePRINT'. The area that you work with is in Brabant and the main target city is Eindhoven. Please keep **thinking aloud** during the execution of the task

1. **Please locate the city of Eindhoven on the map.**
a. Please find the 10 busiest road segments in Eindhoven. Can you find them?
   A: Yes———>go to question (1-b)
   B: No, I cannot find them.———>Please explain why you cannot find them.

   _____

b. On which page / map did you find them?
   _____

c. How well did the application assist you in answering this question?
   A: Very good   B: Good   C: Average   D: Poor   E: Very poor

d. Of these road segments, please find the one where the most cyclists rode their bikes.
e. Can you find it?
   A: Yes
   B: No, I cannot find it. ———>Please explain why you cannot find it.

   _____

f. How well did the application assist you in answering this question?
   A: Very good   B: Good   C: Average   D: Poor   E: Very poor

---

2. Please find all cycling trips whose origin is in Eindhoven and whose destination is in Tilburg
a. Can you find these trips and their origins and destinations?
   A: Yes———>please go to question (2-b)
   B: No, I cannot find them.———>Please explain why you cannot find them.

   _____

b. How many of these routes from Eindhoven to Tilburg are there?

   _____

c. How many trips are made in the morning?
   A:
   _____

   B: No, I cannot get it. ———>Please explain why you cannot find them.

   _____

d. How well did the application assist you in answering this question?
   A: Very good   B: Good   C: Average   D: Poor   E: Very poor

_____

e. From Eindhoven to Tilburg , can you find which routes are chosen mostly?
A: Yes

_____

B: No, I cannot find it. ⟶ Please explain why you cannot find them.

_____

f. How well did the application assist you in answering this question?
A: Very good   B: Good   C: Average   D: Poor   E: Very poor

_____

3. In the urban area of Eindhoven, please find the central train station

a. Can you find it?
A: Yes ⟶ go to question (3-b)
B: No, I cannot find it. ⟶ Please explain why you cannot find it.

_____

b. Can you find all cycling trips that passed the central station and its most adjacent streets?
A: Yes ⟶ go to question (3-c)
B: No, I cannot find them well. ⟶ Please explain why you cannot find them.

_____

c. Of these cycling trips, can you find (1) all trips of which the <u>origin</u> is in this area and (2) all trips of which the <u>destination</u> is in this area?
A: Yes ⟶ go to question (3-d)
B: No, I cannot find it. ⟶ Please explain why you cannot find them and go to (3-f).

_____

d. If you can complete question 3-c, how do you distinguish them?

_____

e. Can you make a comparison of these two spatial patterns of routes? What is the difference?
A: Yes
The difference is:

_____

B: No, I cannot make it. ⟶ Please explain why you cannot make this comparison.

_____

f. How well did the application assist you in answering this question?
A: Very good   B: Good   C: Average   D: Poor   E: Very poor

4. Please find the spatial pattern of all cycling trips in the morning rush hour

a. Can you find it?
A: Yes ⟶ go to question (4-b)
B: No, I cannot do that. ⟶ Please explain why you cannot do it.

_____

b.   On which page / map did you find it?

_____

c.   How well did the application assist you in answering this question?
     A: Very good   B: Good   C: Average   D: Poor   E: Very poor

---

5. Please find all cycling routes on which the average absolute speed is more than 20 km/h

a.   Can you find them?
     A: Yes⎯⎯→please go to question (5-b)
     B: No, I cannot do that.⎯⎯→Please explain why you cannot do it and go to question (5-d)

_____

b.   In which page did you find them?

_____

c.   How many are they?

_____

d.   How well did the application assist you in answering this question?
     A: Very good   B: Good   C: Average   D: Poor   E: Very poor

---

6. Please find the bottlenecks in the cycling infrastructure of Eindhoven

a.   Can you find them?
     A: Yes⎯⎯→go to question (6-b)
     B: No, I cannot find them. ⎯⎯→   Please explain why and go to question (6-e)

_____

b.   On which page / map did you find them?

_____

c.   Of these bottlenecks, can you find 10 where cyclists need to wait for the longest?
     A: Yes⎯⎯→go to question (6-d)
     B: No, I cannot find them. ⎯⎯→Please explain why you cannot find them.

_____

d.   Can you only show these bottlenecks instead of all delay points?
     A: Yes
     B: No, I cannot do that.⎯⎯→Please explain why you cannot do it.

_____

e.   How well did the application assist you in answering this question?
     A: Very good   B: Good   C: Average   D: Poor   E: Very poor

7. Please randomly select one road segment in Eindhoven and select one shown cycling trip randomly.

a. What is the direction of this cycling trip?

A:

_____

B: No, I cannot find that.———→Please explain why you cannot do it and go on.

_____

b. How many road segments were passed by this trip?

A:

_____

B: No, I cannot find that.———→Please explain why you cannot do it and go on.

_____

c. Of these segments, find those segments where cyclists can reach an absolute speed of more than 20 km/h. How many of these road segments are there?

A:

_____

B: No, I cannot do that.———→Please explain why you cannot do it.

_____

d. Can you find what is the average speed?

A: Yes ———→please go to question

B: No, I cannot do that.———→Please explain why you cannot do it and go on.

_____

e. How well did the application assist you in answering this question?

A: Very good   B: Good   C: Average   D: Poor   E: Very poor

APPENDIX 3

Scenario (and self-preparation) of the user research in Amsterdam

The research will take place in Amsterdam and lasts 4 hours (13:00 to 17:00), there are two sessions of this research (1) doing tasks individually, (2) focus group discussion. The second will start when the first finishes. Two employees (Mr.Rutger and Mr.Ewald) will participate the first session and more employees will participate the second one. These is the scenario of the research.

1. Before the testing start

- On interview day, before departing for interview place in Amsterdam, check required documents and devices: hard copies of instruction and task forms, hard copy of this scenario, the laptop with software 'Camtasia'.

-Arrive at the tasting place at 13:00 on 28th Feb 2017.

- When arrive at he tasting place in Amsterdam, get in touch with the liaison man in there and discuss the details about the interview according to the situation including: testing place, testing persons and the order, etc.

- Go to the testing room they provide and start preparing documents and devices:

    Put the hard copies of documents mentioned above on the Table.

    Turn on the laptop and connect it to the Internet.

    Open the software 'Camtasia' and test that it works well.

    Open the website 'BikePRINT' and minimize it.

-According to the order, invite Mr.Rutger or Mr.Ewald to go to the testing room.

- It will start from 13:30

2. Test person welcome and introduction (13:30):

- (To participant) Thank you for coming and welcome, Please take a seat.

- (To participant) Firstly, this is the hard copy of the instruction and task form. Please have a look.

- (To participant) Secondly, I will introduce some key points about contents of this user research on the hard copy and about the test.

- Time for the participant to look at the instruction and the task form.

- (To participant) Do you have any questions?

- (To participant) Whole procedure of testing will be recorded by video&audio&screen recording software 'Camtasia'. This is for academic purposes and will not be released to others. Thanks for your understanding.

-(To participant) if you are prepared, please go to the task part in the document and tell me. We can start

the testing.

3. During the test:

- (To participant) You are the *th one to participate in the test.

- (To participant) During the procedure, please speak aloud whatever you are looking for, seeing,

considering about and feeling. THINK OUT LOUD (this will be recorded).

-Pressing starting button of 'Camtasia' and maximizing 'BikePRINT'.

- (To participant) Now the recording is starting. We can start the testing now based on the list of tasks and

questions.

- (To participant) when you think you find the answer, please use the pen to write down the answer on the

right place on the task form.

- (To participant) Do you have any question so far? Is everything clear to proceed like this? Don't hesitate

to ask me questions anytime. I will help you, but not help you to find the answer to the questions.

- (To participant) Please say 'starting test' and start when you are ready.

- (To participant) When you finish all tasks, please say 'Finished'.

- Observe the whole procedure quietly.

- When the participant finishes the testing, turn down all recording devices and collect the

document.

- (To participant) Thanks for participating. I will now invite the next participant to do the tasks. Please

have a rest. Later on, there will be a group discussion.

- Invite the next one to do the test following the same flow.

- The second one will start at 14:30 and finish at 15:30

4. After the test (15:30---16:00):

- Cost a short time to make a summary of the results of answer sheets.

16:00---17:00

- Invite all participants to be together to have group discussion under the recording.

For the plan of the group discussion, please move to the document 'plan of the group discussion'.

APPENDIX 4

Plan of the group discussion

Step 1 - Discussing Advantages and disadvantages of BikePRINT (20 min)

In this step, all participants are free to tell everything about the testing. For example, they can propose which tasks are hard to answer by 'BikePRINT' and components of 'BikePRINT' they are not satisfied with. Meanwhile, pages of BikePRINT should be projected to screen so we can demonstrate and discuss in a more flexible way. In this session, I will ask some questions to guide them.

Questions related to this step could be:

1. Which is the most difficult tasks for you? Or which is the tasks that is most complicated to deal with?

2. For the difficult task, how do you think that it is difficult to deal with or it cannot be tackled?

3. Which visual representation of visualization(s) in 'Bikeprint' is/are good or bad? Why?

4. What do you like most about the application?

5. How do you think that 'Bikeprint' visualize the Fiststelweek data?

6. What do you not like about the application?

Step 2 - Discussion of analytical questions (20 min)

As I got many their interested analytical from my provided list, it is necessary to have a deep discussion to find what are real wishes and needs behind these analytical questions.

The first part: discuss the marked questions

The second part: acquire other questions and wishes

Step 3 - discussion of the alternative visualization solutions (20 min)

Then I will present basic representations of trajectory and according to the limitations and their wishes to provide some examples of improved visualization solution for them to give comment and preference.

In this step, some questions could be proposed to participants like:

    -Do you like this solution?

    -What do you think about it?

    -What is your suggestion to improve it?

    -and so on.