Towards a Dynamic Isochrone Map: Adding Spatiotemporal Traffic and Population Data

Joris van den Berg¹* (jorisvdberg@live.nl), Barend Köbben² (b.j.kobben@utwente.nl), Sander van der Drift³ (svddrift@dat.nl), Luc Wismans³ ⁴ (lwismans@dat.nl)

¹ UU – Utrecht University, Faculty of Geosciences (Utrecht, The Netherlands)
² ITC – University of Twente, Faculty of Geo-Information Science and Earth Observation (Enschede, The Netherlands)
³ DAT.Mobility (Deventer, The Netherlands)
⁴ CTS – University of Twente, Engineering Technology Faculty (Enschede, The Netherlands)

Abstract. This research combines spatiotemporal traffic and population distribution data in a dynamic isochrone map. To analyze the number of people who have access to a given area or location within a given time, two spatiotemporal variations should ideally be taken into account: (1) variation in travel time, which tend to differ throughout the day as a result of changing traffic conditions, and (2) variation in the location of people, as a result of travel. Typically, accessibility research includes neither one or only variation in travel time. Until recently, we lacked insight in where people were located throughout the day. However, as a result of new data sources like GSM data, the opportunity arises to investigate how variation in traffic conditions and variation in people’s location influences accessibility through space and time. The novelty of this research lies in the combination of spatiotemporal traffic data and spatiotemporal population distribution data presented in a dynamic isochrone web map. A case study is used for the development of this isochrone map. Users can dynamically analyze the areas and people who can reach various home interior stores in the Netherlands within a given time, taking into account traffic conditions and the location of people throughout the day.

Keywords. Isochrone map, dynamic, spatiotemporal, traffic, population distribution, GSM, accessibility
1. Introduction

Accessibility is an interdisciplinary concept that has been used increasingly in a variety of disciplines and studies (Li et al. 2011, Cascetta et al. 2016). In this research, accessibility refers to passive accessibility which is described as 'the ease with which an activity can be reached by potential users in the study area' (Cascetta et al. 2016 p. 45). Accessibility is a spatiotemporal phenomenon, meaning that it changes through space over time (Andrienko et al. 2013). On the one hand, travel times increase and, as a result, the area that can effectively be reached within a given time changes throughout the day. And on the other hand, in order to determine the number of potential users in a study area, you need to know the location of these people through time.

One way of analyzing accessibility is through isochrone maps. Such a map displays isochrones which are the points, lines or areas that can be reached from a given location, within a given time (Bauer et al. 2008, Efentakis et al. 2013, Marciuska & Gamper 2010). Besides visualizing points, lines or areas that are within reach from a given location within a given time, the number of people within an area can be determined by combining the isochrone area with population distribution data. This determines the number of people that could theoretically reach or be reached from a given place within a given time.

A common problem with contemporary isochrone maps is that static travel speeds are assumed when calculating isochrones. As we argued before, traffic conditions, and therefore accessibility, changes significantly over space and time (Li et al. 2011). Using static travel times in accessibility studies and isochrone maps mean that significant variations in accessibility through time and space would be ignored. Moreover, earlier research determining the number of people in isochrone areas (Efentakis et al. 2013, Innerebner et al. 2013) fall short on one crucial point: Spatiotemporal variation in population distribution and movement of people is completely absent.

Currently, mobile data, being GPS tracking, mobile phones and locational media worldwide, provide new opportunities for research into the movement of individuals, the dynamics of traffic and population distribution (Zook et al. 2015). By combining isochrone maps with mobile data, a whole new range of interesting questions can be answered and a more accurate insight into accessibility can be achieved. These mobile data can be used to determine travel times and population distribution data for determining the number of people within areas at a given time. Also, combining spatiotemporal traffic and population distribution data allows knowing where and at what times many people are at the same location, which allows businesses to adjust their opening hours, schedule of events and optimal location

Published as:

The rise of alternative ways to track the movement of people, and the spatiotemporal distribution of populations is particularly interesting in accessibility studies, and until recently has not been used (Järv et al. 2016). The relative newness of these data sources means that no best-practices have been developed yet (Zook et al. 2015).

To dynamically visualize isochrone maps incorporated with spatiotemporal data means a more efficient visualization of spatiotemporal change in accessibility (Innerebner et al. 2013). Besides, it provides a possibility to display interactive statistics allowing easier interpretation of the presented results. As Ullah and Kraak (2015) mention, there is a need for interactive geovisual analytical representation of the produced spatiotemporal data in order to produce useful insights and to make sense out of the data.

Despite these benefits spatiotemporal data might have when implemented in isochrone maps, this implementation can cause new problems. Problems both technically, how to calculate isochrones using vast amounts of spatiotemporal data, as well as how to visualize dynamics in isochrone maps. More data does not necessarily mean more accurate or better results. Ironically, more (spatiotemporal) data means more complications and more effort to conduct useful research (Zook et al. 2015). Although a lot of effort has been put in developing visualization methods that meet the needs to analyze and understand spatiotemporal data (Zeng et al. 2014), options that effectively deal with temporal data in cartography still have not been developed sufficiently (Andrienko et al. 2010, Li & Kraak 2008).

This research aims to tackle the problems mentioned, by combining spatiotemporal traffic and population distribution data in a dynamic isochrone map. The main question in this research is: How can spatiotemporal traffic and population distribution data be incorporated in a dynamic isochrone map? First, we will briefly discuss related work in Section 2 before continuing with the methodology in Section 3, where we discuss how we combined the data and visualized it in a dynamic isochrone map. In Section 4 we discuss the results and end with a discussion and conclusion in Section 5.

2. Related Work

The strength of isochrone maps is to visualize accessibility (Doling 1979, O'Sullivan et al. 2000). Still isochrone maps have been used infrequently in the literature and are often absent from well-known studies on accessibility (O'Sullivan et al. 2000). First applications use static (i.e. fixed) spatiotemporal information on travel times and location of people (mostly based on census data). Assuming static travel speeds and static population distribution has various consequences: users of isochrone maps can make (investment) decisions or interpretations based on overly simplified or erroneous
images of the realities of accessibility (Tenkanen et al. 2016). Isochrone map users, like urban planners, would carry a risk of over- or underestimating accessibility or the number of people within reach in peak hours. Social equity is another related field where problems could occur when using static traffic data (Li et al. 2011, Shaw 2006). People who live relatively close to facilities but suffer from traffic congestion have more difficulties accessing certain facilities than others who are not experiencing traffic congestion. This is especially true in urban areas (Melhorado et al. 2016). Errors could also occur in non-residential areas which are crowded, like airports or business areas. Because officially no one is registered to live in these areas, using static population distribution data in accessibility studies would assume that no one is present in those areas, whereas in the real world significant numbers of people travel to these places. Using spatiotemporal traffic data already proved to be successful in several accessibility studies (Jariyasunant et al. 2010, Jihua et al. 2013, Innerebner et al. 2013, Li et al. 2011, Marciuska & Gamper 2010).

Problems associated with using static travel speeds when calculating isochrones are identified in different studies (Miller et al. 2009, Shaw 2006). These studies conclude that using static travel time in accessibility studies and isochrone maps mean that significant variations in accessibility through time and space would be ignored. Traditional work focused particularly on space (locational) constraints whereas time constraints have been mostly disregarded (Li et al. 2011). A proposed solution is to use spatiotemporal traffic data to calculate isochrones as done by Lee et al. (2009). While research on spatiotemporal traffic data has gained attention in route computation research, spatiotemporal traffic data for calculating isochrones have not received the same consideration (Baum et al. 2015). Efentakis et al. (2013) presented one of the few studies that used static and dynamic traffic data to research differences between the two. They concluded that spatiotemporal traffic data have a ‘huge’ impact on informed business intelligence decisions and showed that the number of potential customers varied between the twenty and forty percent depending on traffic, these variations are quite significant and could be even more significant when taking into account spatiotemporal population distribution. Jihua et al. (2013) created accessibility profiles to display variations in accessibility throughout the day. By plotting the isochrone area in square kilometers versus different hours a day, a better insight in the accessibility of a location was realized. Although the dynamics in travel time also results in variation of people who can reach a location within a certain travel time threshold, there is little to no research in which the dynamics in people’s location is taken into consideration as well.
3. Methodology

Figure 1 shows the methodology developed in this research to construct a dynamic isochrone map. First, a network dataset needs to be prepared (1). Secondly, we calculate the isochrone network using a database network extension (2). Using the isochrone network calculated in the previous step, we can calculate isochrone areas using a specific buffer (3), this is discussed in Section 3.3. The next step is to prepare the population distribution dataset (4). We then combine the isochrone areas with the population distribution data to determine the number of people within the isochrone areas (5). This results in a series of images that display areas, and estimated numbers of people within these isochrone areas (6). The remaining steps consist of visualizing the data processed and calculated in previous steps (7). After adding interactive elements (8), the isochrone map is tested (9) and the visualization process is repeated if necessary.

3.1. HERE network

This research uses HERE Traffic Patterns data (HERE, 2017). HERE Traffic Patterns offers extensive average traffic speed data for 83 different countries, including the Netherlands. These patterns are constructed based on billions of Floating Car observations. The HERE traffic patterns data used in this research contains average driving speeds for every 15 minutes for
every road in the Dutch road network. These speed patterns do not exceed the maximum allowed speed on that specific road. This is because HERE traffic patterns are mostly used for navigation systems and trip planners, and these should not encourage users to exceed the maximum allowed speed. These data are stored and processed in a PostgreSQL database.

3.2. Driving Distance

To calculate an isochrone network, the pgRouting function pgr_drivingdistance is used. This function calculates, starting from a given point, all nodes and edges in the network that have costs less than or equal to a given cost. In our case, costs are given in time. A starting point is entered and the function calculates in all possible directions how far the network can be traversed within a given cost. We used a case study to test our methodology. First, we determined the points from which to calculate isochrones, being the coordinates of different home furniture stores. For illustrational purposes, we used the location of IKEA stores within the Netherlands. These stores served as starting points for the driving distance calculation. Since the driving distance function requires a node on the network as input, the IKEA store locations (retrieved from Google Maps 2016) are snapped to the closest network node.

The driving distance function requires a field that represents the cost and reverse cost per road segment. Since the average speeds are stored in the HERE network data and the road length can be calculated, we created a function that calculates the time it takes to traverse a road segment. In our case, the costs required by the PGrouting function are given as time in minutes. The time it takes to traverse a road can be calculated by combining average driven speeds and the road length. Normally, the driving distance function calculates what parts of the network can be reached from a given point. However, in this research, we are interested in how many people can potentially drive to a store. So instead of stores being a start point, we rather want them to be an end point. This is achieved by switching the cost and reverse cost values. This ensures that the calculation uses the costs for roads towards the IKEA store only.

After the driving distance calculation is completed, a table is created. This table is joined with the output table of the original road network table. The result is a table of all roads on which one can reach an IKEA store, containing a geometry field which can be visualized accordingly.

Despite the fact that pgRouting is fit for handling complex routing computations on extensive network datasets, there are two limitations. When calculating driving distances, pgRouting uses data from one input column for costs. Since one column represents one time step, inaccuracies can occur. If for example average speeds from 9 o’clock with a maximum cost of one hour are used in a driving distance calculation, the 9 o’clock data is used...
during the entire calculation. Ideally, the calculation would have a Time Dependent Dynamic Shortest Path algorithm (TDDSP) meaning the cost field used in the calculation changes according to the time already driven.

Another limitation is that the pgr_drivingdistance function by default returns nodes and edges which do not exceed the maximum input cost. Some road segments are relatively long. If for example they have a node at each end of which one exceeds the maximum cost and one does not, only the latter is returned whereas in reality part of the road segment could still have been traveled before exceeding the maximum cost. Obe and Hsu (2017) have identified this problem as well and describe a possible solution called ‘node injection’. Both of these problems cannot be solved with the current default pgRouting functions.

3.3. Variable Distance Buffer
Since the isochrone area is used to calculate the number of people in reach, it is important to draw an isochrone area around the isochrone network which is as accurate as possible. Marciuska & Gamper (2010) discuss a variety of methods, each with different accuracies. We tested a variety of methods as well. The concave hull, convex hull, alpha shape, link- and surface based approach (Marciuska & Gamper, 2010) and different buffers.

![Figure 2. Variable Distance Buffer versus Concave Hull Approach.](image)

However, each of these methods in- or excluded areas that were actually in or out of reach. This means that when intersecting the calculated isochrone areas with the number of people, large under- or overestimations could occur. We developed a method which in our case is most accurate: the variable distance buffer. Here, a buffer is drawn according to the aggregated costs of the isochrone network. In other words: time already spent can be
taken into account when drawing a buffer. The further driving time increases, the smaller the buffer size. Figure 2 illustrates this solution. The `pgr_driving_distance` calculation explained earlier, stores the aggregated time for each road segment. The variable `distance_buffer` is calculated using a simple IF-THEN statement. If the amount of driven time exceeds a certain value, the buffer size is adjusted.

3.4. GSM Data

This research uses GSM data as a way to measure population distribution. This passive mobile positioning data is aggregated and provided by Mezuro. It originates from Call Detail Records (CDR) collected by a single network provider which facilitates between 30-40% of the Dutch mobile phone usage. This means that by accessing these data we can derive travel information of about one-third of the total Dutch population. No other data source is known that gives travel information on a national scale at a level this high. This data is preprocessed using a validated rule-based algorithm to approximate and classify the number of people within areas based on phone activity for every hour of the day. Because of the known issues regarding the spatial accuracy of determining the location using CDR data (e.g. described in Bonnel et al. 2015), the location data is aggregated at the level of villages. As a result, the Netherlands is split into 1,261 areas for which the number of people is made available, where each city or village is a separate zone. This area definition is the result of earlier analysis by Mezuro of the dataset (CDR and cell tower plan properties) provided. The largest cities in the Netherlands are split into city districts. Mezuro uses a complex algorithm to translate the sample into total estimated number of people, classified in different groups, within these areas. This algorithm takes into account different factors such as the number of people within a GSM area subscribed to the network provider, the number of active phones of subscribers per area and the number of inhabitants per area.

GSM data is privacy sensitive. In theory, it is possible to track someone’s movements. To secure the privacy of mobile phone owners, the CDR’s are anonymized. This means that data of individual phones remain with the network operator. Also, it is not possible to track or filter individual phones out of the provided data. At least 16 phones have to be in the same area before they are registered in the final dataset (Meppelink et al. 2015). As a result of the area definition and by means of aggregating data over multiple days (i.e. all days within a certain month), the impact of this “rule of 16” is minimized (i.e. analysis shows that the impact of this rule is less than 1%). Although the algorithms are tuned for the accuracy of location determination and bias of the available sample, there possibly still remains a bias in the determined number of people in areas.
The aggregated GSM data, from here on simply referred to as GSM data, consists of two tables. The first table contains the administrative areas used. This table does not contain any information on the number of people in that area yet. It merely serves as a spatial reference to the GSM areas. The second table holds the estimated numbers, including further classified population groups in areas for a given month. These two tables are linked through matching area ID’s.

The population groups within the GSM data are classified using observations from the given month. Assumed inhabitants for example, are people that have been observed in the same area during most of the nights in the given month. Regular visitors are observed in a specific area at least 10 times a month. There is a possibility that these people visit these areas because of their job or school. Frequent visitors are people which are observed 3 to 9 times in a GSM area per month. It is hard to determine the goal with which frequent visitors travel to certain areas. It could, for example, be visiting friends or families once a week. The same goes for incidental visitors; People which are observed one or two times a month in a GSM area.

For this specific research, we are interested in the number of people which are located in an isochrone area at a specific time during the day. A differentiation is made only between inhabitants and visitors within the isochrone areas. To determine the number of inhabitants and visitors within isochrone areas at a specific time of day, different calculations were used for inhabitants and visitors. This is because we can further increase spatial accuracy for inhabitants using PC6 points. PC6 points are the centroids of all postal code areas in the Netherlands containing the (static) number of residents in that postal code zone. These points can be used to more accurately determine the distribution of inhabitants within GSM areas. Since PC6 points only contain static information on inhabitants, the method cannot be applied for visitors.

The method for inhabitants intersects the PC6 points with the GSM areas to determine the static total number of inhabitants within a GSM area \((p_e)\). By dividing the number of static inhabitants for each individual PC6 point \((p_{pe})\) by the total static number of inhabitants in the GSM area the fraction of the total number of inhabitants per PC6 point is determined. Using these fractions the estimated inhabitants located in the GSM areas are distributed more accurately. Assuming that this distribution remains the same through time, we can multiply the fraction of each PC6 point which is located in the isochrone area with the associated GSM area dynamic number of measure inhabitants \((I_{are})\) which results in the dynamic number of inhabitants present per PC6 point. The sum of the dynamic inhabitants for these PC6
points (i.e. located within the isochrones area) is the total dynamic number of inhabitants in the isochrone area ($I_{DT}$).

$$I_{DT} = \sum \left( \frac{p_{p}}{p_{t}} \times I_{gsm} \right)$$

$I_{DT}$ = Total number of dynamic inhabitants in isochrone area.

$p_{p}$ = Population PC6 Point.

$p_{t}$ = Total PC6 population GSM area.

$I_{gsm}$ = Dynamic number of inhabitants in isochrone per GSM area.

The method used for calculating inhabitants using PC6 points cannot be applied to visitors since their exact location within the GSM areas is unknown. An alternative approach is to use the isochrone and GSM areas’ surface area. First, isochrone areas are intersected with the GSM areas. The share of the isochrone area ($A_{iso}$) within the GSM area ($A_{gsm}$) can be used to calculate the relative number of visitors in that particular area ($V_{gsm}$). The major assumption in this method is that the visitors are distributed evenly throughout the GSM area. For example, when 50% of the isochrone area intersects with the GSM area, we assume 50% of the total visitors in the GSM area are in the isochrone area. By summing up the visitors within each share of the isochrone area, the total number of visitors in the isochrone area is determined ($V_{T}$).

$$V_{T} = \sum \left( \frac{A_{iso}}{A_{gsm}} \times V_{gsm} \right)$$

$V_{T}$ = Total number of dynamic Visitors.

$A_{iso}$ = Area isochrone within GSM area

$A_{gsm}$ = Total area GSM area

$V_{gsm}$ = Dynamic visitors GSM area
3.5. Visualization

The two major components which are visualized in the web map are the isochrones and the population distribution data. Using the QGIS TimeManager (QGIS TimeManager 2017) isochrone areas were visualized based on the time attribute in our data. The TimeManager filters out, visualizes and exports specific times in the data as images. Using this function, 96 static maps for every 15 minutes of the selected day are exported. Using JavaScript, we created a web page with an interactive interface, to control the playback of the animation.

We have chosen to visualize the population distribution data as an animated line graph, to allow the user to see trends throughout the day. The number of inhabitants, visitors and a total number of people were visualized as a line graph using Microsoft Excel, animated using Microsoft PowerPoint and then added as a video to the web page, synchronized to the map animation using JavaScript code.

We deliberately chose to visualize the population distribution data and the isochrone areas separately to maintain the simplicity of an isochrone map. An alternative would, for example, be to visualize the number of people in a third dimension. We argue that this would overcomplicate the map thereby making it less useful.

4. Results

The resulting Ikea isochrone web map (Figure 3) can be seen online at http://kartoweb.itc.nl/students/iscochroneswebmap/nederland.html.1 It displays an animation of accessibility throughout the day. Users can ‘slide’ through time using the provided time slider and can see the change in area size during different hours of the day. They can zoom into a specific IKEA store to also include the dynamic number of people within the area throughout the day.

As we can see in Figure 3, the difference in the number of people that can reach IKEA Utrecht changes quite noticeably throughout the day. The number of inhabitants (orange line) decrease in the early morning while the number of visitors (green line) increase. This can potentially be explained by the fact that people start commuting. This pattern happens during the end of the afternoon in reverse. Around 16:00, commuters start heading back home. We observe significant drops in the total number of people (black line) around the typical rush hours (06:00-09:00 and 16:00-19:00).

1 If the website animation does not run smoothly, we advise using the step-by-step button.
The variation of the number of people who can reach this location (black line) shows that differences are possible of over 1 million people, roughly resulting in a variation of 30%. Such large differences are certainly relevant for determination of service areas, e.g. choosing the best location for opening a new store.

Another interesting result is that the number of people that can reach an IKEA store located in regions not as much affected by congestion stays relatively stable. Although not visible in Figure 4, but well visible once animated, you can see that the isochrone areas around IKEA Groningen change less in size compared to the isochrones areas around busier IKEA stores such as Utrecht and Amsterdam. For IKEA Groningen we see a relatively stable black line representing the total number of people (Figure 4) which is mainly the results of number of inhabitants and visitors compensating each other resulting in a relatively stable number of people within the isochrone areas. This could mean that the influence of traffic (i.e. travel time variation) has a larger impact on the total number of people that can reach an IKEA store than the movement of people through time.

Figure 3. Isochrones Webmap IKEA Utrecht.
All in all, the differences in the area size and the number of people which can reach an IKEA store within given times justifies the use and need for spatiotemporal traffic and population distribution data in accessibility studies. With differences of up to 1 million people, it would be a mistake to use static traffic and population distribution data.

5. Conclusion and Discussion

We have proven that spatiotemporal traffic and population distribution data can be combined in a dynamic isochrone map to research accessibility. The method used in this research can be used for similar cases without the need to redevelop the methodology. The combination of spatiotemporal traffic and population distribution data is particularly interesting for calculating dynamic service areas which can be used in different fields. A specifically interesting potential use would be the optimization of potential locations for new stores or facilities. One point of attention: some of the calculations were rather time-consuming. Calculating all isochrone areas during a day for the Netherlands lasted approximately 2 days. This should be taken into account when planning to calculate even bigger isochrone areas.

The results of this research are promising although some points can be improved in future research. First of all, pgRouting by default is more focused on the use of static input data for network calculations. Time Dependent Dynamic Shortest Path Algorithms (TDDSP) and a method to increase the accuracy of the nodes returned by pgRouting, for example, would increase the overall accuracy of the isochrone calculations and thus the calculated
number of people within these areas. These functionalities can be added by editing the default pgRouting functions. Also, the interactivity currently offered in the isochrones web map could be extended. It would be nice if users could search for their own address and pan and/or zoom the map.

Besides these methodological points, new, more accurate spatiotemporal data sets might become available in the near future. It would be interesting to calculate isochrone areas using actual traffic speeds of one single day instead of the Traffic Patterns used in this research. Also, more research into GSM data should be conducted to draw better conclusions on the accuracy and usability of these data.

Moreover, the dynamic isochrone map should be tested with actual end-users in future research to evaluate the usability and other potential benefits or shortcomings compared to traditional static isochrone maps. In this research, we claim that adding spatiotemporal dynamics to isochrone maps lead to a better and more accurate insight in accessibility but the potential need and use for such an application are not researched. Another interesting application would be to analyze day-to-day dynamics. This research only focused on a single day but there are significant differences in accessibility between different days as well.

Although there is always room for improvement, especially regarding the visualization of the results, we hope our work encourages further research into dynamic isochrone maps using spatiotemporal traffic and population distribution data. Besides potentially improving the methodology presented in this research, we hope to see relevant new case-studies in which the benefits of a dynamic isochrone map, as presented in this research, are shown.
References


