Discussion paper

Stop and Go detection in GPS-position data

Lars Killaars
Barry Schouten
Ole Mussmann

July 2019

Abstract
Mapping the movement of the general population is important for the development of mobility related policy. CBS collected the time-location data of approximately one week of almost 700 respondents for this purpose. Determining the start and end point of each track is a first step in determining specific points of interest and to find out how a track starts or ends. Certain rules (stop-and-go parameters) are necessary in order to classify the location data in 'stops' and 'tracks'. This paper explores the impact of these parameters on the number of stops, the type of stop and the distance travelled. Results indicate that the statistics are relatively robust for the randomly chosen parameters given to the respondents, but that further reducing the parameters would lead to a significant increase in the number of stops.

1. Introduction

Mapping the movement of people is important to gain knowledge of the mobility of a society and is crucial to determine bottlenecks or vital elements of the infrastructure. Moreover, it provides the necessary information to build prediction models to study the effects of demographic and economic changes on the infrastructure (De Lange et al., 2017). One instrument in measuring travel and mobility is a survey. In the Netherlands, a general population surveys is conducted every year. The current name of the survey is ODiN, which is short for Onderweg in Nederland in Dutch.

The CBS has conducted research on mobility of the Dutch population (CBS 2011-2018) in close collaboration with the Dutch Ministry of Traffic and Infrastructure. The two institutions established an innovation program for the ODiN survey. Up to 2017, ODiN was a mixed-mode survey using a mix of web, telephone and face-to-face interviews. Currently, ODiN is an online survey inviting respondents to provide information on all their travels for a randomly selected day of the week. The required information consists of the locations of all stops, the travelled distances, the means of transportation and any co-travellers.

In 2017 and 2018, CBS developed a cross-platform app (for the source code, see https://gitlab.com/tabi/tabi-app), that tracks time-location when respondents consent to participate. It was anticipated that the survey burden to respondents could be lowered substantially by these automated location measurements. The app (under the name CBS verplaatsingen) was evaluated in the autumn of 2018 in a large scale field test for a sample of 1900 persons within the ODiN innovation program.

The main advantage of an app-based approach is that respondents only have to activate the app and give permission to collect the location data. The app itself would then classify the time-location data into 'tracks' or 'stops' and ask the respondent to add motivations and modes of transport to the data. Therefore, even if the respondent did not give the motivation for the stops, the data regarding the stops and tracks is still collected. In the current online survey, the respondents have to provide all the data themselves and may forget stops due to recall problems or may decide that it was too much effort and not provide any information at all. The app-based approach is expected to place less burden on the respondent. Doing so may improve data quality and is expected to increase the participation rate. In the 2018 experiment (De Groot, 2018), around 30% of sampled persons registered and installed the app and around 20% activated the app for at least a week.

A key point in this data collection method is how the time-location data is clustered in stops and tracks. This is done by an algorithm in the app, see McCool, Lugtig and Schouten (2019), which looks at the relation between the time and distance of a specific time-location entry with other entries before and after (further explained in the Methods section). The choice of a certain time and distance threshold is therefore crucial in the determination of the stops and tracks. In order to test the influence of these parameters, all respondents received a random combination of time and distance boundaries, selected from a pre-specified range. The respondents were unaware of the exact parameter settings.

The influence of these stop-and-go parameters will be evaluated in this study. We aim to answer the following research questions:
- What is the relationship between the stop-and-go parameters and the number of stops that are found?
- What kind of stops are found regardless of the stop-and-go parameters and what kind of stops are only found with certain combinations?
- What is the relationship between the quality of the data and the stop-and-go parameters of respondents?

2. Methods

2.1 The CBS Travel App

The time-location data is collected using a travel app (in Dutch referred to as 'CBS Verplaatsingen app') that the CBS has developed for this research. The app automatically registers the time and location of the device (usually a smartphone) as the sampled persons carry it with them. Under the assumption that the sampled persons have their phone with them and the app remains active, the movement of the sampled persons is tracked during the entire period. The stops and tracks of the sampled persons are determined based on their own stop-and-go parameters and are continuously provided to the respondent. Respondents are asked to provide extra information regarding the stops and tracks, such as motives, transport modes, location names etc.

The sampled persons were sampled from the Dutch population with an age of 16 or higher. In total, 1902 persons were asked to join the test. 951 had participated in previous mobility research and 951 were new sampled persons. Of these 1902, 674 have downloaded the app and 517 persons have sent 7 or more days of data. Of all 17,608 stops measured by the app, 65% have a motive provided by the respondent.

2.2 Algorithm

The app splits all the data points in stops and tracks based on the stop-and-go parameters for each respondent. The algorithm is defined as:

Let $p = 1,2, \dots, n$ be respondents to the travel app and $i = 1,2, \dots, I_p$ be the measurements for respondent p. At each measurement a location and time is stored, denoted by $x_{p,i}$ and $t_{p,i}$.
The two stop detection parameters are a radius r and a duration Δ. Stops are detected when a respondent resides within an area of at most the specified radius for at least the specified duration. This definition still leaves room for different decisions regarding the begin and end of a stop, since a respondent may keep on moving within the stop area.
To express distances we use the operator d, i.e. $d(x, \tilde{x})$ is the distance between the locations x and $\tilde{x}$. In our case, we will use the standard Euclidean distance.
We start by defining the beginning of a stop. Suppose at measurement $i - 1$, the respondent is in movement. When will measurement i be the begin of a stop? We perform the following algorithm to decide whether it is.

- Where $\bar{x}_{p,(i-\Delta,i-1)}$ is the mean location of the collection of points from Δ before the current measurement and the previous point.
- If $d(x_{p,i}, \bar{x}_{p,(i-\Delta,i-1)}) \leq r$ continue, else $i = i + 1$ and return to beginning.
- Set $j = 1$

- Repeat $j \coloneqq j + 1$ as long as $t_{p,i+j} - t_{p,i} < \Delta$ and $d(x_{p,i+j}, \bar{x}_{p,(i+j-\Delta,i+j-1)}) \leq r$
- If $t_{p,i+j} - t_{p,i} > \Delta$ and $d(x_{p,i+j}, \bar{x}_{p,(i+j-\Delta,i+j-1)}) \leq r$ both are true, then i is the beginning of a stop.

When the start of a stop is identified, then the next question is how the centre location of the stop is determined when new measurements come in. Suppose measurements $i - k$ to $i - 1$ all are part of a stop and also a new measurement i is considered to be part of a stop, then the centre location is simply the average location of measurements $i - k$ to i. This location is denoted by $c_{p,(i-k,i)}$ and is defined as $c_{p,(i-k,i)} = \frac{1}{k}\sum_{l=0}^{k} x_{p,i-l}$.

Next we move to the end of a stop, i.e. the beginning of a track. Suppose again measurements $i - k$ to $i - 1$ all are part of a stop and the decision to be made is whether measurement i should be added to the set of stop measurements as well. We would like to use the time-reversed counterpart of the beginning of a stop, i.e. by looking to future measurements rather than to past measurements. If measurement i is the end of a stop, then $i - 1$ is the beginning of a stop in the time-reversed viewpoint. Hence, applying the algorithm above leads to a false for measurement i and a true for measurement $i - 1$. This leads to the following algorithm

- Where $\bar{x}_{p,(i-k,i-\Delta)}$ is the mean location of the collection of points from the beginning of the stop up to the point Δ before the current point.
- If $d\left(x_{p,i}, \bar{x}_{p,(i-k,i-\Delta)}\right) > r$ contintue, else $i = i + 1$ and return to the beginning.
- Set $j = 1$
- Repeat $j \coloneqq j + 1$ as long as $t_{p,i+j} - t_{p,i} < \Delta$ and $d\left(x_{p,i+j}, \bar{x}_{p,(i-k,i+j-\Delta)}\right) > r$
- If $t_{p,i+j} - t_{p,i} > \Delta$ and $d\left(x_{p,i+j}, \bar{x}_{p,(i-k,i+j-\Delta)}\right) > r$ both are true then j is the end of a stop.

Key steps of the algorithm are illustrated in Figure 1 and Figure 2.

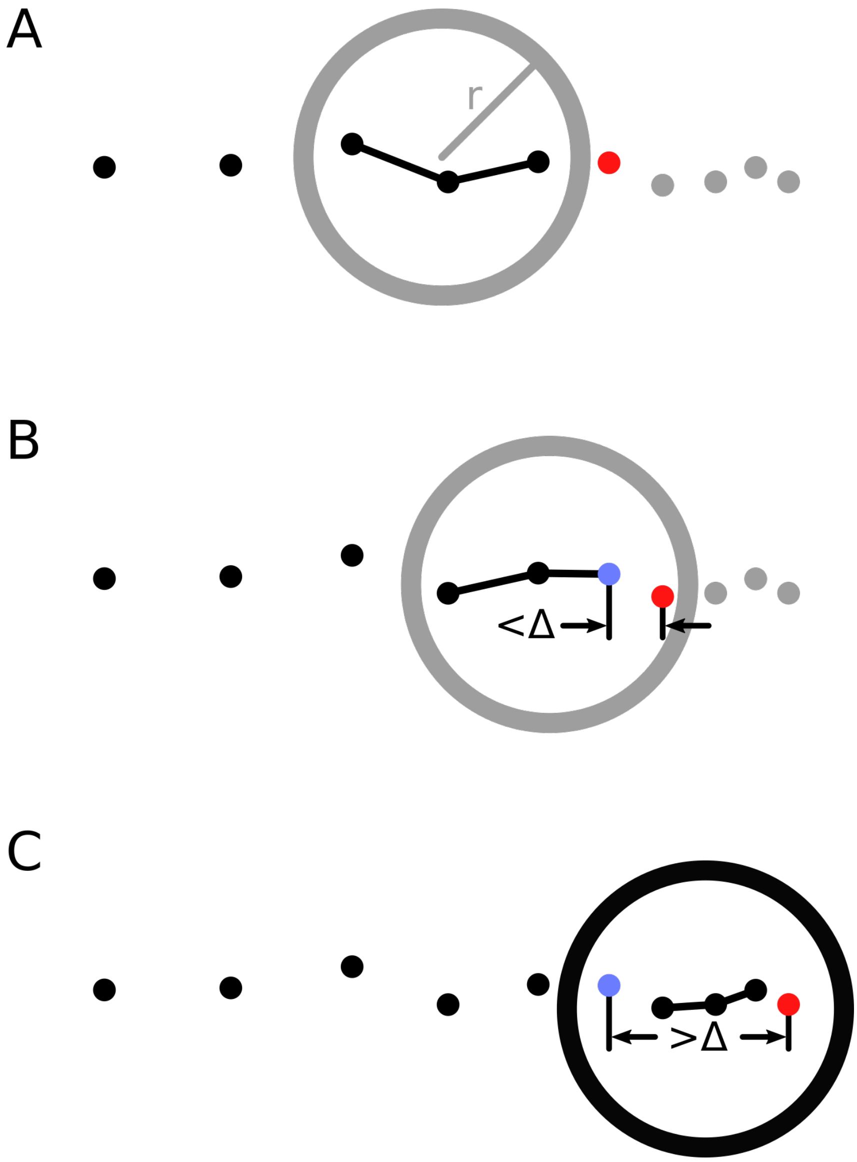


Figure 1: Three time-location entries visualized at the end of a track. The red dot is the current position, the grey dots are future positions while the black and blue dots are positions in previous time-location entries. The circle is the maximum distance (r) from the average location of the connected dots (the reference movement location, or $\overline{x}_{p,(i+j-\Delta,i+j-1)}$). The current position in the first entry (A) is further than r from the reference movement location and the respondent remains therefore in movement. The position in B falls within r from the reference movement location and the previous point is set as time comparison point (blue dot above). The time difference with this point is checked for each following time-location entry as long as the entries remain within r from the reference movement location. If this time difference becomes larger than Δ, the track ends at the blue dot (C). All entries up to and including the current entry form the next reference stop location. In theory, B and C can occur at the same time, if the first entry in r is registered more than Δ after the previous point. If a new point falls outside r, the algorithm switches back to A.

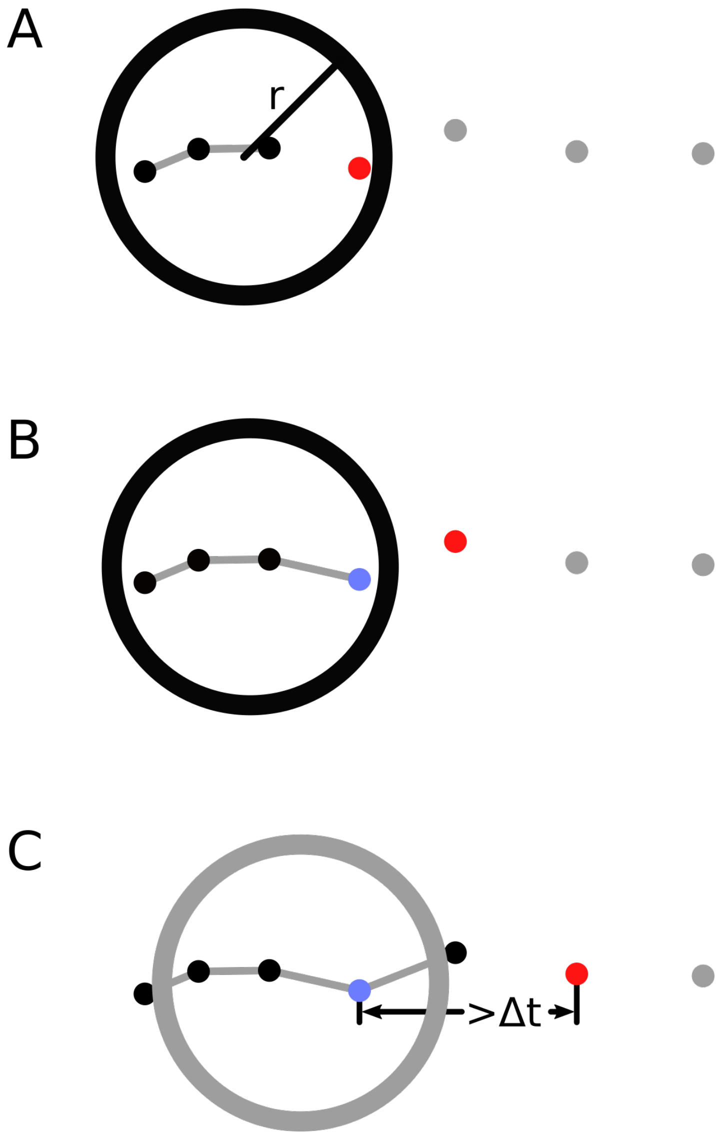


Figure 2: Three time-location entries visualized at the end of a stop. The red dot is the current position, the grey dots are future positions while the black and blue dots are positions in previous time-location entries. The circle is the maximum distance (r) from the average location of the connected dots (the reference stop location or $\overline{x}_{p,(i-k,i+j-\Delta)}$). The current position in the first entry (A) is closer than r to the reference stop location. Therefore, the position is added to the reference stop location and the respondent remains at rest. The position in B falls outside r and the previous point is set as time comparison point (the blue dot). The time difference with this point is checked for each following time-location entry as long as the entries remain outside r from the reference movement location. If this time difference becomes larger than Δ, the stop ends at that location (the red dot in C). All entries up to this entry are added to the reference stop location. In practice, B and C can occur at the same time, if the first entry outside r is registered more than Δ after the previous point. If a new point falls within r, the algorithm switches back to A.

3. Results

In this section, answers are provided to the research questions. An important side remark needs to be made beforehand. The time-location measurements from the mobile device sensors show missing data. See McCool, Lugtig and Schouten (2019) for a detailed explanation of the causes of the missing data. The most prominent cause is instability of the app on the background. Mobile devices operating systems may deactivate an app when other apps are running on the foreground. This can be prevented to a large extent byt redesigning the app and asking the respondent to disable battery saving mode for the app. The missing data imply that sometimes stops or tracks are missed. Furthermore, travelled distances may be underestimated for these trajectories as a shortest distance is computed rather than the actual travelled distance.

3.1 What is the relationship between the stop-and-go parameters and the number of stops that are found?

In the field experiment, the time and radius parameters were varied and randomly allocated to sample units. The time parameter was randomly selected from the set {2 , 3 , 4 , 5} minutes and the radius parameter was randomly drawn from the interval [60, 100m]. In preliminary analyses for this paper it was found that within the chosen parameter sets, the variation in number of stops was relatively small. For this reason, the modified algorithm of section 2.2 was used to simulate numbers of stops outside the range of parameters in the field test. In addition, as a very useful spin-off, for every respondent stop detection was applied for multiple values of the parameters, making the data much richer. Obviously, the respondents only provided data on transport mode and purpose of travel for the specific set of parameters that was assigned to them.

Stop detection was applied for 35 different stop-and-go parameter combinations. The time parameter is varied from 1, 1.5, 2, 3, 5, 7 to 10 minutes and the radius from 10, 30, 50, 70 to 90 meter. The mean number of stops per day for all users can be used to show the relation between the stop-and-go parameters and the number of stops (Figure 3). It shows that the mean number of stops slowly rises when the time and radius parameters are reduced, until the sharp increase at the low end of the scale for both parameters.

Not suprisingly, the largest number of stops are found when a combination of 1 minute and 10 meters is used, giving an average of 14.5 stops per day per person. The amount of stops decreases to 2.6 stops per day for a combination of 10 minutes and 90 meter. At first, the decrease in the number of stops for an increase in radius may be counter-intuitive. The larger the area in which a stop may take place, the more stops may be expected. This is not true; the larger areas actually cluster multiple stops that are only found when applying a smaller radius.

The range of settings actually used in the field test (60-100 meter and 2-5 minutes) indeed has a small gradient and shows an increase from 3 stops per day (90 meter and 5 minutes) to 4.5 stops per day (50 meters and 2 minutes).

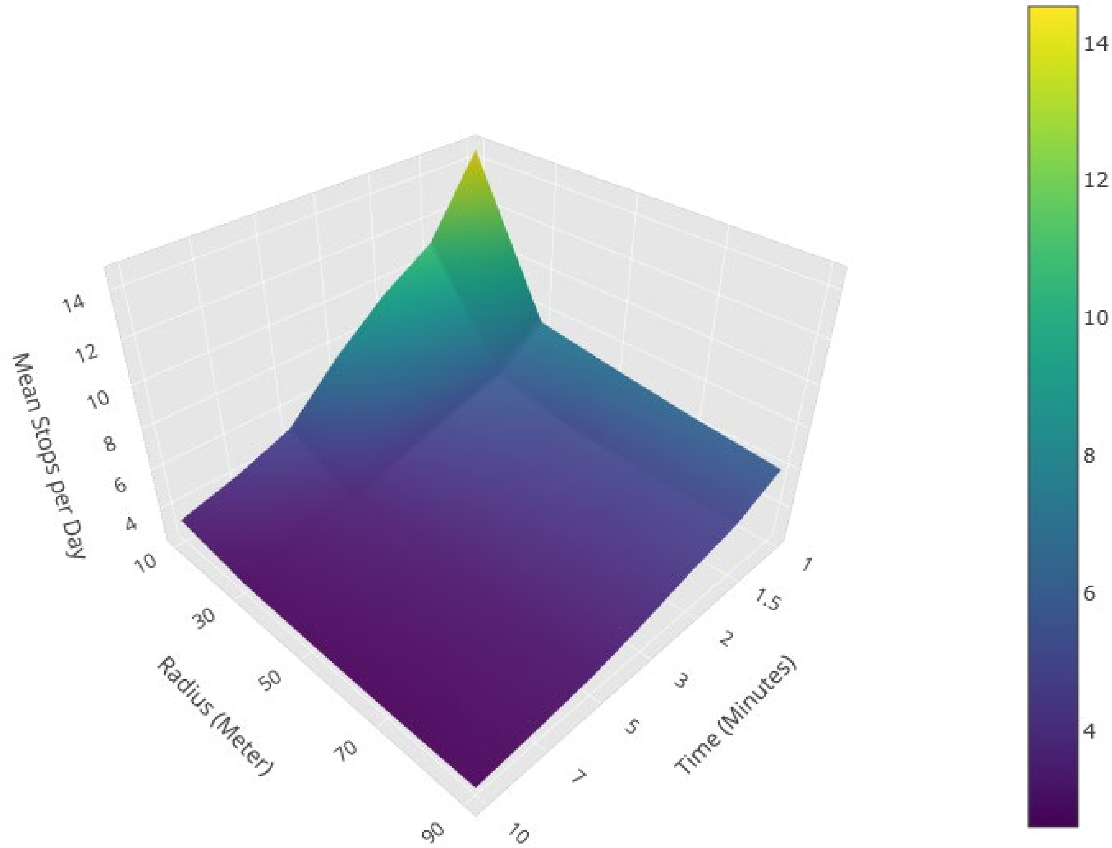


Figure 3: Mean stops per day for all stop-and-go parameters. Each data point contains the average of the mean number of stops per day for that specific stop-and-go parameter combination for all 674 respondents with data.

The variance of the mean number of stops per day has a similar pattern. The variance increases with lower stop-and-go parameters, as some people gain more stops than others. The variance rises especially for parameters below 2 minutes or 30 meters, when the mean number of stops rises fast as well.

3.2 What kind of stops are found regardless of the stop-and-go parameters and what kind of stops are only found with certain combinations?

Different combinations of stop-and-go parameters will result in a different number of stops that are found (Figure 3). Small sections of position entries of a respondent may trigger the stop detection rules with a certain parameter combination, but not with another. Figure 4 shows the location of all stops of one person during one day to illustrate a stereotype pattern with different parameters, but the same location data. 14 stops are made with a parameter combination of 1 minute and 10 meters (Figure 4a), some of which closely overlap with each other and are only visible when zooming in. The number of stops decreases to five (four are visible) and three stops (two are visible) for the other two combinations shown (3 minutes and 50 meters and 10 minutes and 90 meters, Figure 4b and c). The app itself presented the 5 stops in Figure 4d to the respondent, who gave the motives 'Visit' (2x), 'Get or bring goods', 'Get or bring persons' and 'Home'. The 'Home' stop is located at the bottom right corner, 'GetBringPersons' at the top and the 'Visits' (twice at the same location) and 'GetBringGoods' are the leftmost stops. The combination 3 minutes and 50 meters detects the same stops as the app, but the combination 1 minute and 10 meters does not find any new stops at different locations. The stops detected with the other combinations are split in smaller pieces as the stop detection rules are triggered by small movements 'during a stop'. In this example, one of the visits is split in two and the final 'Home' stop is split in 9 smaller stops. Increasing the parameters (Figure 4c) actually removes stops, as the 'Get or bring goods' and 'Get or bring persons' are not distinct enough to be seen by these stop-detection rules.

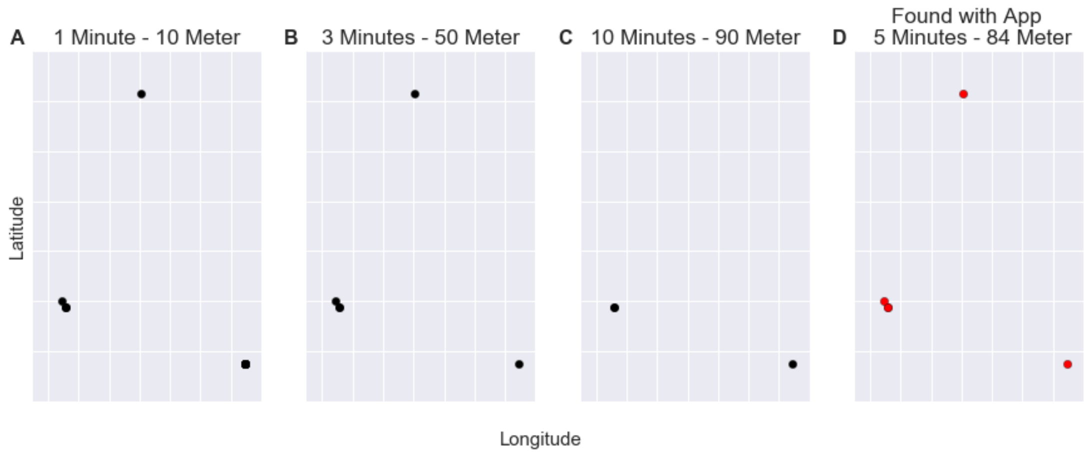


Figure 4: Anonymous stop locations for one day of one user found by different parameter combinations. A-C show the found stops with simulated parameters, while D shows the stops found during the field test and were presented to the user.

This effect occurs for most users and for all changes in parameters (Table 1). Even though the results are skewed because some long stops are always found (e.g. sleep, phone remains at the same place, app is inactive for longer period), shorter stops and tracks occur when the parameters are decreased. The standard deviation is large due to these long stops and tracks. The fact that the averages can decrease despite this high variability, indicates that mostly short stops and tracks are created. This is also shown in Figure 5. Stops and tracks with a duration shorter than 10 minutes are almost non-existent with a stop-and-go parameter of 5 minutes or larger, but form the majority of tracks and stops with shorter parameters. This shift is less pronounced for the radius (not shown in this paper), but the trend remains towards shorter tracks and stops with a smaller radius parameter.

More and shorter stops/tracks also means that the track distance becomes smaller (Table 1). The mean distance is 9.2 km for a time parameter of 1 minute and even 6.8 km for a radius parameter of 10 meters. Figure 5c shows that the fraction of tracks with a length below 1 km increases from 50% to 80% for both parameters, while the fraction of tracks longer than 1 km is reduced. The number of stops/tracks increases for all categories when the time or radius parameter is reduced (not shown in this paper), but the proportional increase of the smallest categories is larger than the rest, leading to the increase shown in Figure 5.

The estimated durations for the stops and tracks in table 1 are surprisingly long. This is the consequence of the missing time-location data from the mobile device sensors. The stop detection misses stops within tracks and tracks within stops, depending on the context. As a result, both tend to be somewhat larger. This overestimation can be adjusted for to some extent, but this is outside the scope of this paper.

Table 1: Mean and standard errors of the duration and distance of stops and tracks for each time and radius parameter used.

Time (minutes	Stop duration (minutes)	Track duration (minutes)	Distance of the track (meters)
1.0	163 ± 0.875	54 ± 0.660	9.248 ± 259
1.5	209 ± 0.948	73 ± 0.786	11.480 ± 282
2.0	238 ± 1.04	83 ± 0.797	12.262 ± 280
3.0	286 ± 1.15	103 ± 0.870	12.176 ± 247
5.0	368 ± 1.30	137 ± 0.996	14.124 ± 250
7.0	428 ± 1.42	161 ± 1.04	15.384 ± 265
10.0	480 ± 1.41	197 ± 1.26	15.693 ± 256

Radius (meters)	Stop duration (minutes)	Track duration (minutes)	Distance of the track (meters)
10	145 ± 0.779	71 ± 0.769	6.798 ± 204
30	290 ± 1.13	104 ± 0.920	12.976 ± 275
50	335 ± 1.24	111 ± 0.913	14.932 ± 293
70	355 ± 1.29	115 ± 0.910	15.796 ± 300
90	368 ± 1.32	116 ± 0.900	16.337 ± 305

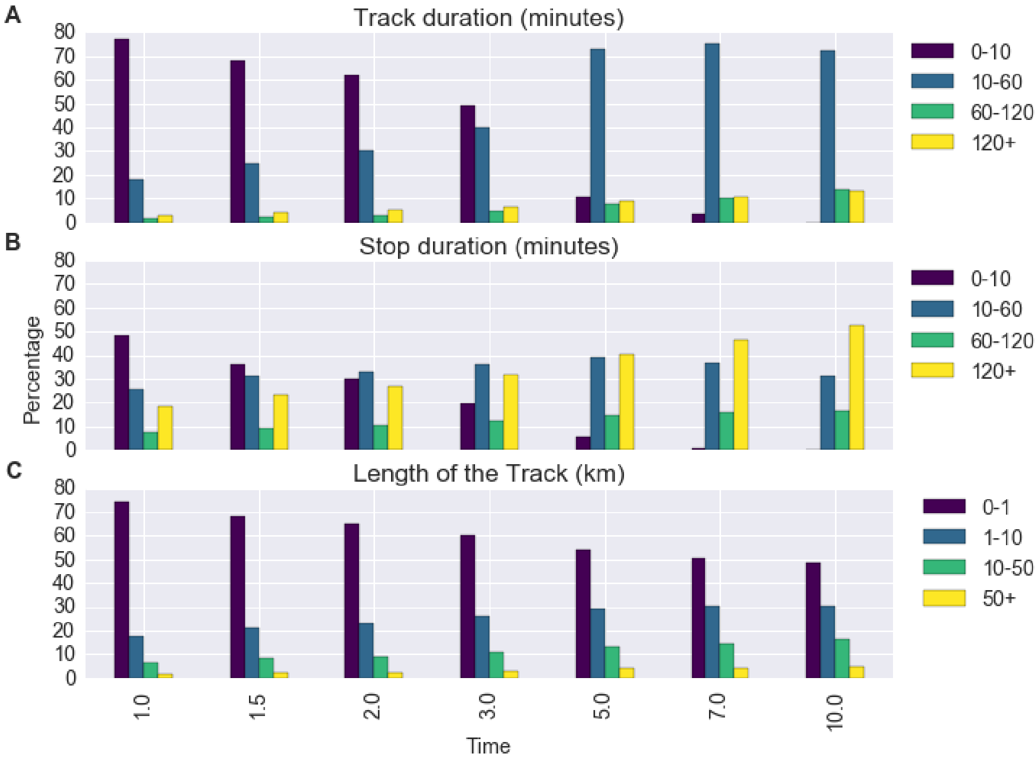


Figure 5: Percentage of stops/tracks with a duration/distance that falls within the specified boundaries on the right. Shown for all used time parameters.

3.3 What is the relationship between the quality of the data and the stop-and-go parameters of respondents?

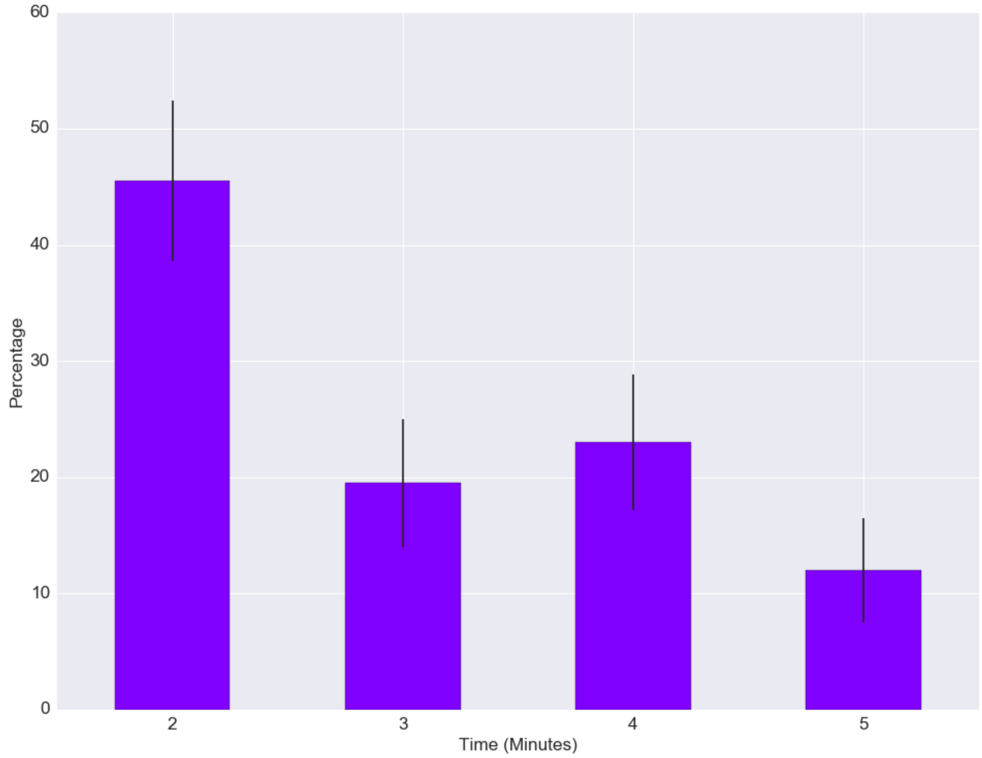
There were multiple instances where respondents could indicate that the app had made errors in detecting stops and/or tracks. Respondents had the opportunity:

1) to label stops as 'Location incorrect' or 'Transfer'
2) to give feedback in questions they had to answer for each calendar day
3) to participate in the evaluation survey after the field test.

Three daily questions had to be answered before a day could be validated and completed: 1) Do you have any comments about the stops and tracks reported for today?, 2) Was there anything unusual about this day's travel? and 3) Did you have your mobile device with you during all travels?

All sample units that registered the app were invited to particpate in an online survey that was conducted three weeks after the end of the app data collection.

The stop labels and answers were used as a first indication of the quality of the data. It must be noted, however, that also respondents themselves may be wrong or apply a subjective definition of a stop or track.



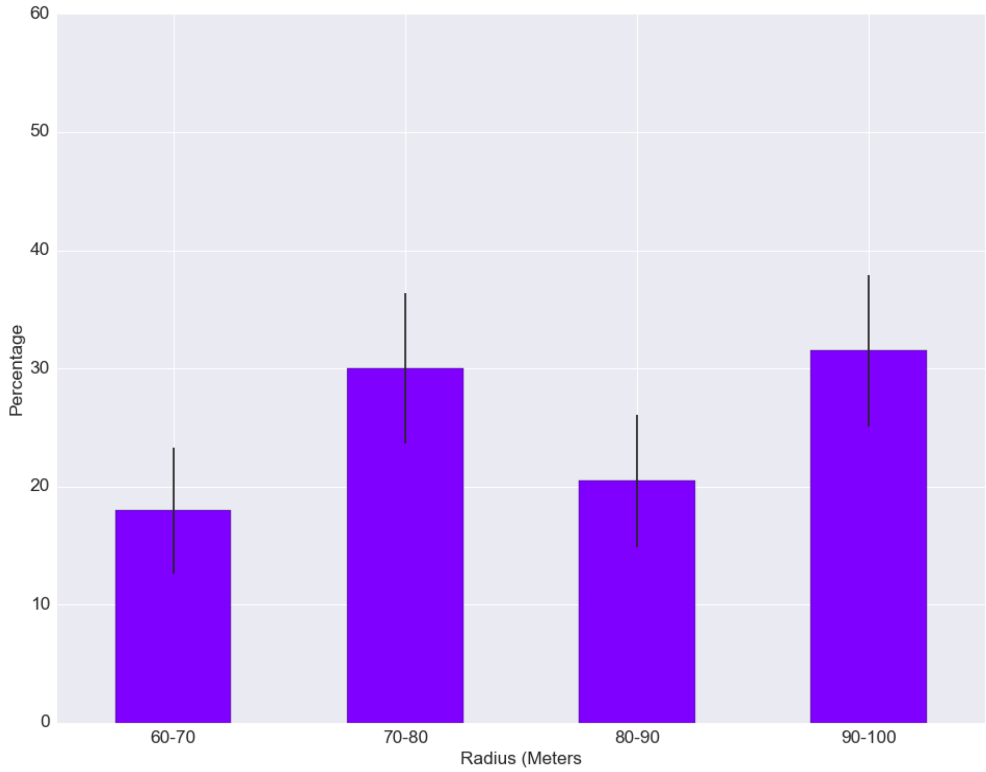


Figure 6: Percentage of stops that had the motive "LocationIncorrect", separated by the stop-and-go parameters of the respondents.

Of all the 17,608 stops registered by the app, 200 were labelled as "Location incorrect" by 95 different users. The distribution of stop-and-go parameters of these users is shown in Figure 6. Almost 50% of the labelled stops come from 31 respondents with a time parameter of 2 minutes, while the other 3 groups have a significantly lower amount of stops labelled "Location incorrect". If other motives are included that could be selected for 'extra' stops, such as public transport transfers or traffic jams, this pattern becomes less clear, but is still significant. The distribution for the radius parameter is more equal for both stops labelled "Location incorrect" and for the collection of motives, here no clear pattern is visible.

A different indication of the quality of the detected stops came from the online evaluation survey for which all persons were invited who had previously registered the app. These persons were asked if the number of stops reported by the app was higher, equal to, or lower than the number of stops that they thought they had made in the data collection period. Regardless of their stop-and-go parameters, the majority (60%) answered that the app reported less stops than they thought they had made (Figure 7). The group answering that the app found more stops than they expected was always the smallest group. It is unclear what kind of stops were missed by the app, as there was no option for the respondent to create stops themselves and no further questions were asked to determine the nature of stops that were missed. This issue was not mentioned in other feedback from the respondents. It must be stressed that the app itself was unstable and sometimes stopped working. The app only registered the location when it was active, so this issue caused gaps in the coverage for respondents. Since it is unclear what kind of stops are missed, this could be due to the stability of the app. Further evaluation of this issue falls outside the scope of this research. Related to this was the question in the evaluation survey if the app had registered a stop or track that was not made. The results of this question were similar to the results in Figure 7.

Other indicators of data quality were used, apart from respondents direct feedback. The amount of days that users participated can show if the app was working correctly and was not taking too much effort from the respondent. There was no clear relationship between the stop-and-go parameters and the amount of days that respondents participated.

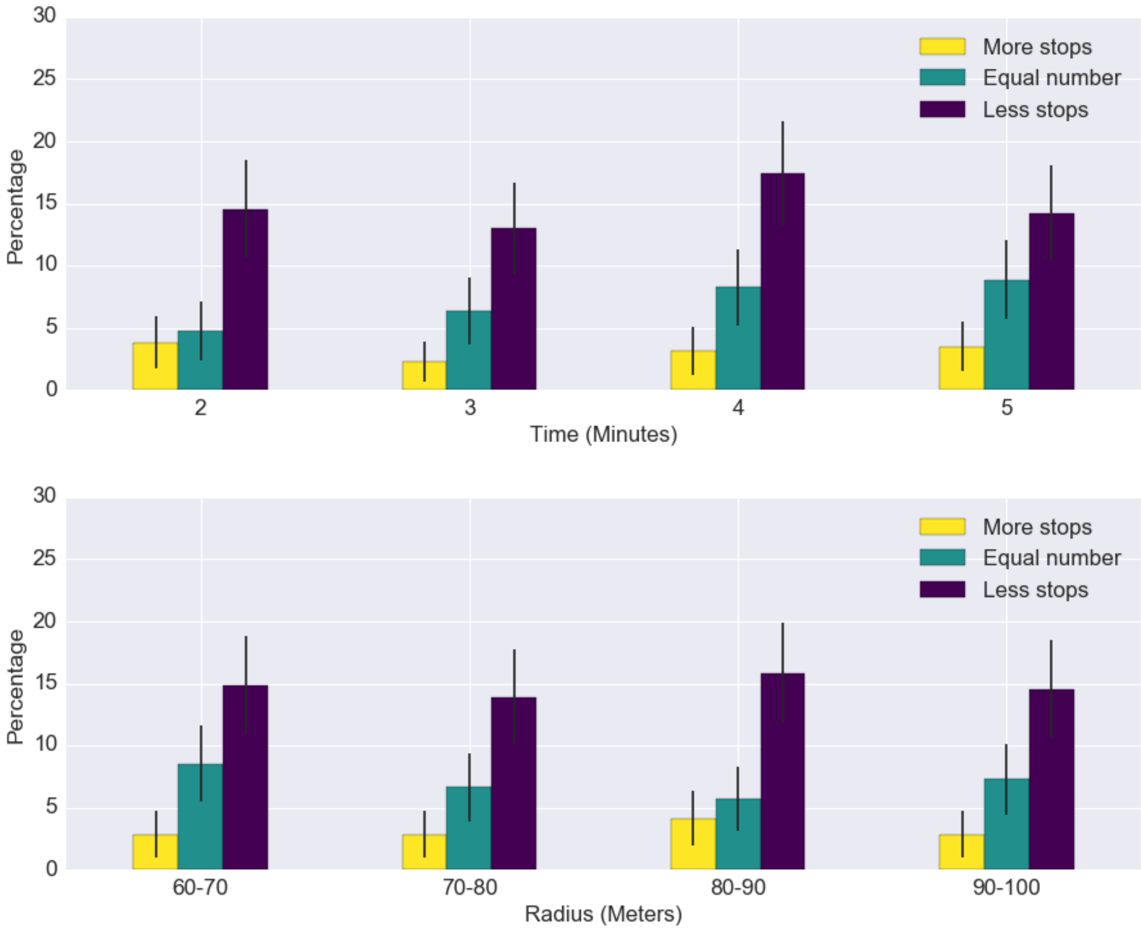


Figure 7: Answers of respondents to the question: How does the number of stops that the app found compare to the amount of stops you expected?

4. Discussion

The results indicate that there is a sharp increase in the number of stops when the stop-and-go parameters are set below 2 minutes and 30 meters. Since the users of the app had a time parameter of 2, 3, 4 or 5 minutes and a radius between 60 and 100 meters, well above the threshold, the data quality of the detected stops is not impacted significantly by the stop-and-go parameters given to the respondents.

4.1 What is the relationship between the stop-and-go parameters and the number of stops that are found?

There is a clear relationship between the number of stops and the stop-and-go parameters. Lower parameters mean that more clusters of location points will trigger the stop-detection rules and will eventually detect all stops that the respondent has actually made. However, it will also yield extra stops by splitting a group of location points that might just trigger the stop-detection rules but are no real stop of the respondent at all. On the other hand, increasing the parameters will lead to the disappearance of stops as some clusters of location points will not trigger the stop-detection rules anymore. This appearance and disappearance of stops (from Figure 4a to b to c) happens between all parameter combinations and does not make a distinction among stops that are intentional or stops that are the result of short delays like traffic lights or traffic jams. A stop is created if the stop-detection rules are triggered, regardless of the circumstances. This is the main drawback of the approach used in the travel app. Given the used parameter combination for the respondents, it is highly likely that the shortest stops are not found. This is strengthened by the fact that most respondents mentioned that the app did not pick up all stops.

4.2 What is the relationship between the quality of the data and the stop-and-go parameters of respondents?

The fact that the given parameters mostly fall in the relatively flat section of the 'number of stops detected' curve (Figure 3) is related to the fact that there was no significant relationship found between the data quality and the respondents. Only users with a time parameter of 2 minutes have a significantly higher proportion of stops labeled 'Location incorrect' than other users, but no other relationships were found.

4.3 What kind of stops are found regardless of the stop-and-go parameters and what kind of stops are only found with certain combinations?

Certain kinds of stops are found regardless of the parameter settings. These robust stops are clearly defined by a duration longer than 10 minutes and have all their location points grouped close together. They are characterized by motives such as 'Work', 'Home' or 'Visit' and are usually activities that last a longer amount of time and have a fixed location. As such, they are always detected by the algorithm since they trigger the stop-detection for all used parameters. More dynamic activities are harder to detect by all parameters, as their duration might be shorter than 10 minutes and use a larger area. Examples of these activities are bringing/picking up persons or objects, walking your dog or parking your car before doing the actual activity. Therefore, different parameters will result in different detected stops as the specific data point that triggers the stop-detection rules, if any, will differ for the different parameter combinations.

Similar results were found in a study by Zhao et al. (2015). They determine the stops of a user in multiple steps, where their first steps are comparable to our method of only using a time and radius parameter to determine the stop location. They use a time parameter of 1 minute and a radius of 50 meters to find all possible stops of a user. The found stops are afterwards selected or merged based on context information about the location, motion during the stop or time-location difference with other stops to account for GPS errors or measurement difficulties in general. In an example they find 106 candidate stops with 5 spatially distinct locations on a specific day for a user, but actually report 3 stops at 3 locations after their data cleaning steps. We observe similar patterns, where reducing the stop-and-go parameters increases the amount of stops, but not necessarily the amount of distinct stops. Stops are mostly divided into smaller stops at almost the same location, with some individual stops or small clusters appearing on the route between the large clusters. However, the respondents mention that not all stops are found, so it is important to determine what kind of stops are still missed and if this is due to the stability of the app.

Multiple respondents reported that the app was unstable and shut itself down, especially on iOS devices. This effect is also visible in the data. There are gaps in the location data of several respondents, making it impossible to determine the stops and tracks for these periods. Moreover, when the app is restarted, it is impossible to know if the user is currently moving or not, possibly leading to errors in the first period after such a shutdown. The respondents had no opportunity to add stops themselves, meaning that stretches of time are not accounted for in the research. The extent of this problem remains unclear and falls outside the scope of this study.

Consequently, the respondent should still play an important role in GPS based stop-detection studies, especially if a high accuracy is one of the requirements. Using only time and radius as stop-and-go parameters leaves ambiguity in the results as it is impossible to distinguish short stops from inaccurate GPS readings or travel disturbances based on them alone. Therefore, user input is still required and can take different forms based on the needs of the researcher and the level of input required. Examples are asking verification for each found stop, training a machine learning algorithm for each respondent, using (user specified) context information or using motion detectors etc. However, there is no 'best' solution, as all options require either extra effort from the respondent or more battery/computation power of the device from the respondent, both possibly leading to a higher dropout rate from the respondents. Using only time and radius might not provide the most accurate results, but using a reasonable setting will identify the most important stops and tracks of a respondent for a minimal effort.

5. Conclusion & implications

The main conclusion of this study is that the applied stop-and-go parameters in the travel app do not significantly impact the resulting number and type of stops. The range of parameters given to the respondents (2 - 5 minutes and 60 - 100 meters) was higher than the parameter combinations found to significantly increase the number of stops (2 minutes and 30 meters). Some have a time parameter of 2 minutes, which is on the found limits, but the radius is always large enough to dampen the effect of the low time parameter. As a result, it is not expected that the distribution of the stop-and-go parameters over the respondents of the ODiN research will significantly impact the overall results of this study.

For future studies that only use a time-radius combination for stop detection, multiple options can be chosen, depending on the desired accuracy or respondent burden. If the goal is to obtain all possible stops from the respondent, the parameters should be set low. The time parameter should be set to 1 minute and the radius parameter to 10 meters. This will yield all stops from the respondent, but will also split several of the longer stops, meaning that some post-processing might be required before presenting the stops. However, if a low respondent burden is required with little to no post-processing, a combination of 3 minutes and 60 meters is advised. This is sufficient to stay away from the sharp increase in stops, but small enough to catch smaller stops and tracks.

We experienced improvement in the accuracy of stops when parameters such as the accuracy of the location measurement or dynamic reference locations were also taken into account. Further improvement may come from external information about a location or motion sensors. Merging the stops afterwards is also a strong tool to improve the accuracy of the stops.

6. References

CBS (2011). Onderzoek Verplaatsingen in Nederland 2010. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2012). Onderzoek Verplaatsingen in Nederland 2011. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2013). Onderzoek Verplaatsingen in Nederland 2012. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2014). Onderzoek Verplaatsingen in Nederland 2013. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2015). Onderzoek Verplaatsingen in Nederland 2014. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2016). Onderzoek Verplaatsingen in Nederland 2015. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2017). Onderzoek Verplaatsingen in Nederland 2016. Den Haag. Centraal Bureau voor de Statistiek.

CBS (2018). Onderzoek Verplaatsingen in Nederland 2017. Den Haag. Centraal Bureau voor de Statistiek.

De Groot, J. (2018). Onderzoeksdesign Aanschrijving verplaatingen app, evaluatie verplaatsingen app. Den Haag. Centraal Bureau voor de Statistiek.

De Lange, M., Huibregtse, O. (2017). Verkenning alternatieve inwintechnieken voor verplaatsingsgegevens. Den Haag. Kennisinstituut voor Mobiliteitsbeleid.

Zhao, F., Ghorpade, A., Pereira, F. C., Zegras, C., & Ben-Akiva, M. (2015). Stop detection in smartphone-based travel surveys. Transportation Research Procedia, 11, 218-226.

Colophon

Publisher
Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress
Statistics Netherlands, CCN Creation and visualisation

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contactform: www.cbsl.nl/information

© Statistics Netherlands, The Hague/Heerlen/Bonaire 2019.
Reproduction is permitted, provided Statistics Netherlands is quoted as the source.