From Location Tracking to Personalized Eco-Feedback: a Framework for Geographic Information Collection, Processing and Visualization to Promote Sustainable Mobility Behaviors

Dominik Bucher, Francesca Mangili, Francesca Cellina, Claudio Bonesana, David Jonietz, Martin Raubal

Abstract

Nowadays, most people carry around a powerful smartphone which is well suited to constantly monitor the location and sometimes even the activity of its user. This makes tracking prevalent and leads to a large number of projects concerned with trajectory data. One area of particular interest is transport and mobility, where data is important for urban planning and smart city-related activities, but can also be used to provide individual users with feedback and suggestions for personal behavior change. As part of a large-scale study based in Switzerland, we use activity tracking data to provide people with eco-feedback on their own mobility patterns and stimulate them to adopt more energy-efficient mobility choices. In this paper we explore the opportunities offered by smartphone based activity tracking, propose a general framework to exploit location data to foster more sustainable mobility behavior, describe the technical solutions chosen and discuss a range of outcomes in terms of user perception and sustainability potential. The presented approach extracts mobility patterns from users’ trajectories, computes credible alternative transport options, and presents the results in a concise and clear way. The resulting eco-feedback helps people to understand their mobility choices, discover the most non-ecological parts of their travel behavior, and explore feasible alternatives.

Keywords: Mobility Tracking, Mode Detection, Trajectory Clustering, Eco-feedback, Sustainability

Acknowledgements. This research was supported by the Swiss National Science Foundation (SNF) within NRP 71 “Managing energy consumption” and is part of the Swiss Competence Center for Energy Research SCCER Mobility of the Swiss Innovation Agency Innosuisse.
1. Introduction

Many smartphone apps perform automatic mobility and activity tracking, with increasing precision both
in path recognition as well as in the identification of transport modes [22]. While many applications focus
on health and fitness aspects [53], or passively track in the background, for example to give location-based
recommendations (e.g., [47]), there has been a recent surge in applications concerned with individual mobility
behavior [17, 19, 25, 6]. Such apps are often driven by ecological incentives such as reducing CO₂ emissions
from travel or generally increasing the sustainability of personal transport. The key components aim at
making people aware of their mobility behavior, i.e., at showing them the impact of their behavior on the
environment, on their financial budget or on the time spent traveling.

Still, in many applications users just check the individual routes they travel and do not get a complete,
yet simple and immediate understanding of their mobility patterns. It is left to the app users to assess which
journeys are suitable for optimization and what the effects of a behavioral change on the environment are.
Getting the right eco-feedback, defined as “feedback on individual or group behaviors with a goal of reducing
environmental impact” [18], and making users aware of their mobility patterns and the consequences they
tail, is acknowledged as a necessary—though not sufficient—condition towards more sustainable mobility
patterns (cf. [17, 16]). For example, current eco-feedback technologies lack the distinction between systematic
and unsystematic travel behavior, the provision of a personalized and meaningful assessment of possible
changes and the presentation of the various effects behavioral changes entail.

In the Swiss-based GoEco! project [10], we addressed these issues, taking advantage of the wide
availability of smartphone devices to encourage people to reduce their use of cars. Using a specifically
developed smartphone app [7, 9], we performed a large scale field test, monitoring the activities of several
hundred volunteer citizens from Southern Switzerland and the City of Zurich. During three distinct mobility
tracking periods distributed over a year, each one lasting at least six weeks, the movements and transport
mode choices of participants were recorded. To assess the effects of eco-feedback on mobility behavior,
no feedback was given to the study participants in the first and last period (i.e., they were only tracked
to collect a baseline dataset). Starting after the first period, they were also stimulated with eco-feedback
and gamified activities, with the aim of persuading them to improve their mobility behavior. While the
gamification elements (such as goal-setting, challenges, badges or leaderboards) were integrated directly into
the smartphone app, eco-feedback was additionally given in the form of a report booklet, sent to the study
participants via e-mail. In this paper, we mainly focus on this non-interactive eco-feedback report which was
sent to the participants at the end of the first period.

This report contained basic elements such as summaries of distances traveled, durations spent on different
modes of transport or energy consumption and CO₂ emitted as effect of the recorded travel behavior.
These summaries were available at various time scales, such as the overall study period or individual weeks.
Additionally, the report contained a section about detected systematic journeys (appearing multiple times
within a few weeks, such as traveling to work or school), highlighting which of them could be optimized with
little loss in travel convenience by adopting a proposed alternative. An estimate of the possible impacts on
the environment (energy consumption and CO₂ emissions) rounded off the report. In post-experimental
surveys we assessed the plausibility of the detected systematic routes and the proposed alternatives. Here, we
present the algorithms used to detect systematic travels and to compute feasible alternatives. In addition, we
provide an estimate what a sample of individuals like the one of the GoEco! project could potentially save in
terms of CO₂ emissions by changing mobility behavior to more sustainable modes of transport, according to
the proposed alternatives.

To the best of our knowledge, this is one of the first experiments of its type. Therefore, it required the
development of a framework for collecting, analyzing and reporting individual mobility patterns and potential
for changes towards more sustainable mobility choices. The proposed framework includes: (i) location
tracking; (ii) identification of transport modes; (iii) assessment of individual mobility patterns (modal share
and mobility footprint); (iv) identification of systematic itineraries; (v) identification of more sustainable

---

1. See [goeco-project.ch](http://goeco-project.ch) for details and recent updates.
mobility choices; (vi) presentation of the outcomes as personalized feedback. Section 2 introduces related work from the domain of mobility behavior assessment and eco-feedback provision. Section 3 presents the proposed framework, structurally following the six parts mentioned. In Section 4 we provide additional details, illustrating the core methods of our framework. The results of the application of the above framework and methods to data gathered as part of this study are presented in Section 5. Finally, we discuss the strengths and weaknesses of our framework and point out remaining open questions in Sections 6 and 7.

2. State of the Art in the Analysis of Activity Data to Understand Mobility Behavior

In the past, several studies have aimed at not only analyzing people’s mobility behavior, but also at making it more environmentally sustainable by using mobile applications and a feedback loop. Examples include UbiGreen [17], Tripzoom [9], MatkaHupi [25], SuperHub [50], PEACOX [41, 33] and Traces [51]. However, these studies often involve small sample sizes (up to several dozens of users), short study periods (up to a few weeks), lack a control group or solely perform qualitative assessments [21, 45]. For example, UbiGreen [17] was one of the first smartphone-based eco-feedback applications and featured a visual representation of the ecological impact of travel behavior in the form of a tree that lost or gained leaves. If one chooses sustainable transport modes on a certain day, the tree would grow greener. The thirteen participants mostly responded in positive ways to the application, but a systematic quantification of the UbiGreen effects was not performed in the paper. Building on these previous studies, our research tries to apply their findings to a large sample over the duration of one year, and studies the adoption of eco-feedback related to mobility and the resulting changes in behavior.

However, before eco-feedback can be provided or sustainable travel alternatives can be proposed, movement resp. mobility needs to be recorded. This is a topic which has attracted interest from various disciplines and has also proven to be useful for numerous applications other than the promotion of sustainable mobility behavior, including transportation planning (e.g., [23]) or personalized routing (e.g., [28]). Commonly, movement is tracked using Global Navigation Satellite System (GNSS) receivers, e.g., the GPS sensor in the mobile phone of a user. Numerous studies have proposed approaches to augment this tracking data, e.g., by inferring the performed activities or the used transport modes by considering additional sensor data (see for instance [55, 43, 44]). Such research has allowed the development of several apps for tracking users’ itineraries and activities [22]. To obtain the results in Section 5 we applied the methods presented within this paper to GPS data tracked by such an app (namely MovesR⃝), which automatically recorded user trajectories and identified visited places and routes. To improve the transport mode recognition, an additional new classifier was developed that learned from previously tracked and validated data and additionally considered spatio-temporal context (cf. [7]).

Once such movement data is available, it must be analyzed and mined for observable and interpretable individual mobility patterns. The most basic of these patterns include modal splits and aggregate values (such as total CO2 emissions or travel duration). Most of the previously mentioned studies used such aggregate values, e.g., in UbiGreen [17], where the daily number of “eco-friendly” journeys is compared to the number of car journeys. In our framework, we provide methods to analyze mobility on a much more detailed (and individual) level by inspecting individual journeys, with the aim of providing users with a more thorough feedback and thus offering more assistance for behavior change. For this task, several trajectory pattern mining techniques can be employed [20]. Previous studies on mobility data mining have particularly focused on the identification of visited places from raw GPS-data by assessing prolonged permanence in a small area of the moving person [35, 29, 3], and the derivation of traveled routes between these locations by map matching techniques [31, 38]. By assessing the similarities between trajectories of several individuals, these methods generally discover knowledge about collective mobility behaviors, like common places of interest, frequently visited locations, the way of moving, etc. This is useful information in urban planning and transportation management for instance [37, 32]. Other works, closer to our aim of generating meaningful eco-feedback,  

2 A fitness tracking app discontinued in July 2018.
use pattern mining techniques to detect individual mobility habits by looking for regularities in personal trajectory data \[54, 36, 27\].

After finding patterns in the trajectory data, personal feedback and proposed sustainable travel alternatives can be provided. Typically, this requires a personalized, multi-modal routing application. When computing optimized routes, personalization is often achieved by incorporating individual preferences and constraints as parameters of route segment traversal costs into the routing algorithm (e.g., \[2, 30\]), or using a preceding heuristic \[8\]. A range of techniques has been developed especially for computing routes on large-scale, dynamic multi-modal transportation networks (e.g., \[5, 13, 12\]). An appropriate aggregation of the output of such planning systems can then be given to the user as feedback on possible more sustainable alternatives.

Giving meaningful eco-feedback is a well-researched topic (cf. \[48\]). Commonly, one tries to satisfy psychological (autonomy, competence and relatedness) and social needs (achievement, affiliation and intimacy, and leadership and followship) and should respect the current context of a user (e.g., how motivated a person is, if a user has the ability to perform a certain behavior, if the timing of the feedback is right, and if a user is in the correct stage of a behavioral change process) \[48\]. In terms of eco-feedback particularly for mobility behavior, there are many more open questions, as mobility is highly individual and to a large degree depending on personal, temporal and environmental context \[49\]. As it is generally difficult to passively collect data on a user’s context (e.g., if the user needs to carry heavy luggage for a certain journey), in looking for available more sustainable alternatives we here primarily respect temporal and environmental context factors, such as the available transport modes at a given date and time.

3. A Framework for the Generation of Eco-Feedback

Generating eco-feedback is a complex procedure involving different steps. Figure 1 shows the individual steps of our framework to induce behavioral change in the context of mobility choices. The system includes a mobility tracker installed on the personal smartphone of each user. The tracker’s task is to automatically and passively register all routes traveled by each participant. The system then assigns a transport mode (TM) to each detected journey, based on a transport mode classification algorithm and/or on direct interrogation of the user (e.g., the recipients of eco-feedback are often given the possibility to validate and adjust a recognized transport mode). The collected routes are then stored in a database and analyzed following three main goals:

- to profile the user mobility patterns, in terms of modal share and related footprint (global analyses);
- to evaluate the potential improvement each user can achieve regarding overall sustainability of her mobility choices;
- to identify similar trajectories (i.e., systematically traveled routes), for whom targeted and realistic suggestions for more sustainable mobility choices (alternatives) can be proposed.

Finally, the outcomes of this threefold analysis (eco-feedback) are presented to the user in a report.

3.1. Mobility Tracking

In our mobility tracking framework, a trackpoint is the most basic object. It consists of a longitude, a latitude, and a timestamp. Trackpoints are aggregated into routes which are defined as transfers from a point of interest (POI) to another, where a POI is a place visited on purpose, to perform or attend some activity (such as working or shopping). Each route represents a process, which has a well-defined start and end point, both in space as well as in time (these start and end points are central for several parts of our method, while the actual trackpoints of the route play a minor role). The distance traveled is the sum of the distances between all trackpoints, and the duration of a route is the difference between the timestamps of the end and start positions. The longitude and latitude coordinates of the start and end points of a route will be referred to as visited positions, whereas the POIs connected by a route are physical places that cannot be described by a single pair of GPS coordinates; indeed, also because of the inaccuracies of the tracking system, different (yet close) visited positions may refer to the same POI.
Each route can be traveled with more than one transport mode. Segments, on the other hand, are parts of a route covered with a single mode of transport. A stay point is any point where a user spends some minutes, e.g., when changing from one transport mode to another. For example, a route traveled by \[\text{walk} \rightarrow \text{train} \rightarrow \text{walk}\] would consist of two walk segments, and one train segment and will include two stay points (i.e., the train stations).

The goal of a tracking app is therefore to collect a thorough yet concise set of trackpoints, and to identify the visited places and stay points which allow to partition the whole set of trackpoints into routes and segments. Common issues in the development of tracking apps for large-scale mobility data collection studies ([39][42][44]) include battery consumption, availability on diverse operating systems, and varying classification accuracies due to different sensors (cf. [7] for an overview). There are several apps that promise off-the-shelf solutions to these problems (with varying success). For example, in our study, the commercial app Moves® was employed, which allowed custom apps to connect and download tracking data through an application programming interface (API). Moves pre-processed the data from mobile sensors and provided trajectories as sequences of trackpoints already partitioned into segments, routes and stay points, which were of sufficient quality for eco-feedback and mobility analyses [7]. While the exact workings of Moves were not disclosed publicly, from our experience it recorded a single trackpoint approx. every 5 minutes (depending on the transport mode and GPS availability), used the accelerometer for transport mode identification and treated places where users spend more than approx. 10 minutes as staypoints. Alternative options to employing Moves would have been to either use a more pricey paid app, integrate an open-source solution, or develop a new tracking app from scratch. As the focus of this study was on the exploration and analysis of an eco-feedback intervention, there were not sufficient resources available for any of the alternatives (especially the development of a reliable tracking app from scratch is a non-trivial task).

3.2. Transport Mode Detection

In order to assess the ecological impact of peoples’ mobility, a transport mode has to be allocated to each segment. Similarly to most commercial trackers currently available, Moves only distinguishes between a small set of transport modes, specifically walk, bike and motorized transport. As a more detailed classification is necessary for ecological footprint analyses, we developed a classifier with the ability to distinguish between thirteen transport modes (car, electric car, motorbike, scooter, train, bus, tram, plane, ship, bicycle, electric bicycle, kick scooter and foot). Details of the algorithm for transport mode classification are given in [7].
To improve the reliability of the assigned transport modes, we present all the routes identified by *Moves* to the users, showing the transport modes assigned by the classifier. The users are then asked to validate the routes, i.e., confirm or change the mode of transport and delete a route if it is “particularly wrong”, either regarding the places visited or the travel time. This possibility to delete routes was mainly put in place as sometimes the recorded routes are completely unrealistic (e.g., on the border of a country, where due to changing mobile phone network carriers, the cell tower positioning method yields bad data) and would greatly distort following data processing. On the one hand, this manual validation supplies a set of labeled data that can be fed back into the algorithm to improve the accuracy of the future transport mode classification, while on the other hand, it ensures a high accuracy of the collected data. As the user validation continued along the whole study period, the dataset of routes and transport modes used as a base for all our computations can be considered as reasonably correct and not affected by the accuracy of the classifier.

3.3. Mobility Patterns: Modal Share and Mobility Footprint

The most basic individual mobility features, which summarize the multitude of data automatically gathered by the app, are the *total distance traveled* and the related *travel duration*, averaged over a specific time period. In this study, we opted for a weekly time period, as it is easily understood by users, it allows including both systematic and non-systematic routes (i.e., routes that are resp. are not traveled several times within a short time frame, such as a week), and accounts for the variability of daily mobility needs arising from individual weekly schedules.

A second set of mobility features includes the individual shares of the different modes of transport used. Our eco-feedback shows this modal share based on the average weekly kilometers traveled by each mode of transport, as this is directly related to energy consumption and CO$_2$ emissions. Of course, it would also be possible to show the modal share based on traveling times or the number of routes, however, these aspects are more relevant when optimizing schedules than for improving ecological footprints.

To simplify the feedback, some transport modes are aggregated, such as *car* (both internal combustion engine and electric car), *public transport* (train, tram and bus), *bicycle* (both conventional and electric), and *foot*, while *other* encompasses all the remaining modes of transport. Finally, the eco-feedback includes information on users’ ecological footprints, expressed in terms of energy consumption (kWh) and CO$_2$ emissions. Both their average values per week and per kilometer traveled are shown to the user. To provide such estimates, we relied on the Mobitool consumption and emission factors [46], which depend on the mode of transport, refer to a single kilometer traveled in Switzerland and take into account the consumption and emissions of the full life-cycle (see Table 1). For many transport modes these factors heavily depend on the involved power generation systems (e.g., power from renewable sources leads to much fewer CO$_2$ emissions than power generated by fossil-fuel power plants). As such, these factors and the resulting energy consumption and CO$_2$ emission values are specific to Switzerland. In order to apply the framework to another region, the power generation systems within that region have to be analyzed and taken into account accordingly.

3.4. Systematic Mobility

A central element of our eco-feedback are the suggestions of mobility alternatives specific to the itineraries traveled by each individual user. Since it is not possible (and neither useful) to suggest alternatives for all ever traveled routes, we focus on the systematic ones, i.e., those traveled multiple times within a certain (short) time frame. Systematic itineraries are of interest, as a behavioral change in these situations is easier, since they can be planned in advance and become part of the daily routines. Also, they have a larger potential to reduce energy consumption and CO$_2$ emissions, as they are repeated over time.

To provide applicable eco-feedback and suggest credible alternative modal choices, our approach moves from individual routes to *loops* of routes starting and ending at the user’s home place. Without this step (i.e., if suggestions of alternatives are based on routes only), the system might end up suggesting unreasonable or conflicting solutions within the same loop. For example, in a car-based $[home \rightarrow workplace \rightarrow city$}
Table 1: Energy consumption and CO₂ emission factors used to estimate a user’s ecological footprint.

<table>
<thead>
<tr>
<th>Transport Mode</th>
<th>Energy Required (kWh/km)</th>
<th>CO₂ Produced (gCO₂/km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>0.91</td>
<td>197.23</td>
</tr>
<tr>
<td>Bus</td>
<td>0.64</td>
<td>145.41</td>
</tr>
<tr>
<td>Tram</td>
<td>0.55</td>
<td>37.47</td>
</tr>
<tr>
<td>Train</td>
<td>0.14</td>
<td>7.32</td>
</tr>
<tr>
<td>Bicycle</td>
<td>0.04</td>
<td>7.64</td>
</tr>
<tr>
<td>Walk</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

centre → workplace → home loop, suggesting to replace the car-based route [workplace → city centre → workplace] with a bicycle-based one is not plausible, since it cannot be assumed that a bicycle is available at the user’s workplace. Or, in a car-based loop [home → workplace → supermarket → home], suggesting public transport to reach the workplace might not be acceptable, if soon after work the user directly goes to a large supermarket in the suburbs, which is not served by public transport. To avoid such inconsistencies, we need to explicitly take into account the constraints that choosing a means of transport on a specific route imposes on the following routes. In conclusion, by considering clusters of systematic loops, we can account for systematic mobility patterns of users and at the same time identify reasonable alternatives.

The specific algorithm adopted for the identification of systematic and non-systematic loops is described in Section 4.1. Instead of blindly searching for clusters of similar itineraries based on standard trajectory clustering algorithms, we focus on the needs and the limitations of this study and develop a more interpretable algorithm by answering to some questions such as “how should a loop be defined?”, “how different can loops be, which belong to the same cluster?”, “how many times and how frequently has a loop to be traveled, in order to be considered systematic?”. Answering these questions, we arrived at our definition of systematic loops and at the selection of heuristics adopted by the algorithm, by keeping in mind the final goals of the intervention, i.e., to suggest feasible and meaningful alternatives that can effectively induce a behavioral change, and to assign realistic values to the individual potentials for change. The developed algorithm also accounts for the limitations and inaccuracies of the tracking system and tries to correct them.

3.5. Alternatives and Potential for Change

The assessment of someone’s potential for change is primarily based on the comparison to optimal travel behavior, where optimal in this case is computed with respect to CO₂ emissions. While many factors could be included in this computation (such as possibilities to telecommute, or the rescheduling of activities), here we only consider the individual routes and the modes of transport chosen. This choice is primarily made based on the available data, which is not detailed enough to infer accurate schedules and contexts of users. We consider the contribution of all the systematic loops and all other routes (i.e., routes for which no systematic loop could be constructed) to the overall CO₂ production and, for each of them, we identify an optimal alternative solution, i.e., a different itinerary and modal choice that allows reaching all the POIs visited in the original loop or route, while reducing the overall CO₂ production. Note that this means that we consider the alternatives for all traveled journeys to compute the optimal mobility behavior. However, in the final feedback, suggestions for concrete alternatives are only given for routes that are frequently traveled. The potential for change is then computed as the difference between the actual emissions and those that would have been achieved by replacing all loops in each cluster of systematic loops and all other routes by the corresponding optimal alternative. When no alternative is found, the original loops or routes are retained.

Therefore, given any route consisting of a number of segments annotated with a mode of transport, we compute the CO₂ produced as the sum of all the segment lengths multiplied by an emission factor which depends on the mode of transport. Again, we refer to the Mobitool values already introduced in Table 1 specific to Switzerland. Duration, distance, and start and end points of routes and segments are defined as in Section 3.1.
A range of different optimization criteria could have been used instead of CO\textsubscript{2} emissions, such as the total covered distance, the duration spent traveling, or the required energy. We focus on CO\textsubscript{2} production as we are concerned about giving eco-feedback to people (which ultimately should result in a reduction of greenhouse gases emissions), and this has the benefit of reducing the required energy as well, as there is a strong correlation between energy consumption and CO\textsubscript{2} emissions.

Concerning the identification of alternative solutions, in order to successfully replace a route, an alternative has to respect the following requirements:

1. It cannot use a mode of transport which is not available to the user (e.g., if someone does not own a bicycle, alternatives containing bike segments are invalid). While this can be extracted from the validated routes themselves, the study participants were also initially asked which modes of transport they have access to.
2. Its duration should not be excessively long (e.g., if traveling by public transport takes a user three times as long as traveling by car, she will most likely ignore the suggestion). This is less problematic for short routes, as there the metric of interest is usually the additional time an alternative takes (e.g., replacing a 10 minutes car route with a 20 minutes bicycle route is acceptable, even though it increases the traveling time by 100%).
3. Its CO\textsubscript{2} production should be substantially lower than the original one (e.g., more than 5% lower, to account for tracking and positioning inaccuracies). This is usually the case when switching to a more eco-friendly mode of transport.

Requirement 1 is particularly important when considering loops, especially those including car and bicycle transport modes. When starting a loop with one of these modes of transport, we need to make sure it is brought back to the starting location (i.e., the user’s home place). On the other hand, if the first route does not start with any of these two modes, they cannot be used later on in the loop either. While the first requirement imposes a strict criterion for removing unacceptable alternatives, the other two form an optimization problem. The algorithm to identify alternative solutions for routes and loops that meets the requirements above is described in Section 4.2.

3.6. Report

The analyses described above lead to an eco-feedback which is presented to each user in the form of a short report. The report is divided into two parts: the first part analyzes the potential for change when only systematic loops are replaced by the proposed alternatives, while the second part includes the complete set of loops and routes recorded for each user. As changes induced in systematic itineraries may have a larger impact on the overall energy savings and may be more easily undertaken and maintained by the user, the first part of the eco-feedback details both the identified loops, and the suggested alternatives. For each systematic loop recognized by our algorithm, users receive a feedback similar to the one depicted in Figure 2. On the left hand side, a representative loop, selected among the original loops of the systematic cluster, is shown on the map. In the example of Figure 2, the representative loop consists of three POIs and starts with a short walk, followed by tram segments. On the right hand side, the suggested alternative is presented. It is suggested to take the bicycle, which is not only more eco-friendly, but even faster in this case. It produces 0.18 kg CO\textsubscript{2} each time the loop is traveled and can be completed in 48 minutes, while the original modal choice produces 0.28 kg CO\textsubscript{2} and takes 1 hour.

In addition to these individual loop assessments, the second part of the report provides aggregate results about the overall travel duration and CO\textsubscript{2} emissions that could have been achieved if the user had always adopted the optimal alternative suggested by the system. An example of the summary displaying the total potential for change of a user is shown in Figure 3. In this example, a user, originally choosing the car for 93.9% of her routes, can see that she could potentially reduce its use to 32.6%, resulting in average CO\textsubscript{2} savings of 56.7 kg per week (from 78.5 to 21.8 kg/CO\textsubscript{2}/week). In a similar way, on average the energy consumption could be reduced by 238.6 kWh per week (from 369.6 to 131 kWh/week).

While the feedback about systematic loops shows concrete suggestions for improvement, the overall potential for change (Fig. 3) is a less tangible indication. For this reason, in the second part of the report we
4. Methods

In this section we provide more details about the methods to identify loops and systematic loops and to suggest valid, realistic and more sustainable alternatives for each user mobility choices.

4.1. Identification of Systematic Loops

In a first step, individual trajectories must be aggregated into loops, which then can be clustered into systematic loops, based on the frequency in which they appear.

**Loops.** To identify loops, we first introduce the concept of connected routes. Two consecutive routes are connected if the first one ends at the same location where the second one starts (here we are assuming that routes are chronologically ordered). In practice, the GPS coordinates of the end position of a route

---

**Figure 2:** View of the alternative for a systematic loop as displayed in the eco-feedback report. On the left is the original loop, on the right the alternative.

**Figure 3:** The modal choice of a single user, as identified based on six weeks of monitoring using a tracking app (A), and the modal choice suggested in the provided eco-feedback (B). Next to each figure, the average weekly traveled distances are shown.

<table>
<thead>
<tr>
<th>Distances [km]</th>
<th>Car</th>
<th>Bicycle</th>
<th>Pub. Trans.</th>
<th>Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>396.1</td>
<td>15.6</td>
<td>0.8</td>
<td>8.9</td>
</tr>
<tr>
<td>Total</td>
<td>421.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
seldom coincide with those of the start position of the subsequent route, and their distance can even be rather large, due to the (occasionally substantial) inaccuracies produced by the activity tracking tool used in this project. Therefore, we consider as the same place any two positions whose distance is less than 500 m. For the loop identification, all routes of a user are considered.

We begin from the first route that starts at the user’s home place (which is known from a study pre-survey) and follow the sequence of connected routes until we find a route that ends at the user’s home. It can happen that the ending point of a route (hereafter referred to as route 1) and the starting point of the next one (hereafter referred to as route 2) have a distance larger than 500 m, for instance due to GPS malfunctioning. In this situation, one could artificially connect the two consecutive unconnected routes by adding an unobserved route going from the end point of route 1 to the start point of route 2. However, we preferred not to add unobserved routes, partially because the cause of such a situation is not necessarily the loss of a route, but it may also be an error in the GPS positioning system that localizes the user in a different position than she actually is, and opted for deleting unconnected routes. To this purpose, we used the following algorithm.

Case 1: If route 2 starts at the end point of a previous route of the loop under construction, we remove the intermediate routes and continue the construction of the loop.

Case 2: If the starting point of route 2 has never been reached before in the loop under construction, route 2 is discarded and the construction of the loop continues with the next route.

To better understand this algorithm, consider the following example (see Figure 4).

Let us assume that user X has visited the sequence of points \(\text{home, work, restaurant, work, home}\).

Case 1: If only the \([\text{home} \rightarrow \text{work}], [\text{work} \rightarrow \text{restaurant}]\) and \([\text{work} \rightarrow \text{home}]\) routes are recorded by the activity tracker, we will have a route ending at \(\text{restaurant}\) (the second one) followed by one starting at \(\text{work}\) (the third one), and it would not be possible to complete the loop. However, we notice that \(\text{work}\) has already been visited before, and thus, by removing the \([\text{work} \rightarrow \text{restaurant}]\) route, we complete the \([\text{home} \rightarrow \text{work} \rightarrow \text{home}]\) loop. The identified loop is just a subsection of the correct one, yet it has been actually traveled and should be taken into account.

Case 2: If, instead, the activity tracker missed the \([\text{work} \rightarrow \text{restaurant}]\) route, the second, i.e., \([\text{restaurant} \rightarrow \text{work}]\) route will start from a position which was not seen before in the loop under construction. Then, the \([\text{restaurant} \rightarrow \text{work}]\) route is removed and the construction of the loop continues with the third \([\text{work} \rightarrow \text{home}]\) route, which is connected to \([\text{home} \rightarrow \text{work}]\). Once again, the subset \([\text{home} \rightarrow \text{work} \rightarrow \text{home}]\) of the original loop is obtained.

![Figure 4: Two possible ways a loop of routes is mistakenly recorded by the activity tracker.](image)

**Systematic Loops.** Once all loops have been reconstructed, to identify clusters of *systematic loops* we look for groups of loops that are similar. To compare loops, we represent them as sequences of POIs in chronological order. POIs are identified by analyzing the geographical distribution of the visited places. As already mentioned, several different (yet close) visited places may correspond to the same POI, not only because of the spatial extent of the POI but also because of the inaccuracies of the GPS and tracking systems.
Therefore, it is necessary to define a threshold distance $D$ below which different GPS coordinates are assumed to represent the same POI. Once again, the threshold $D = 500$ m is adopted. Such a choice implies a low resolution in distinguishing two close POIs. Clearly, this is an application specific setting that should depend both on the scope of the analysis and the precision and accuracy of the collected data.

The aggregation of similar items (in our case, geographical positions of visited places) is typically done using a clustering algorithm. As the number of clusters is unknown in this case, we use hierarchical clustering [24], which allows building the set of clusters by setting a maximum acceptable distance between two items belonging to the same cluster. The POIs of a user are obtained by applying hierarchical clustering to the set of her $n$ visited places, i.e., the start and end points of all the routes included in her loops. Hierarchical clustering builds a hierarchy of clusters by starting from $n$ clusters containing a single place each and aggregating at each next level of the hierarchy the two clusters of the previous level which are closest, based on a specified dissimilarity measure. We apply the average linkage distance strategy (cf. [40]) that defines the dissimilarity between two clusters as the average distance between pairs of elements (one in each cluster). The aggregation process is stopped when the distance between the two closest clusters exceeds the predefined threshold of $D = 500$ m. Finally, each POI is identified by the centers of mass of the corresponding cluster of visited places.

Each loop can then be translated into an ordered sequence of POIs, so that loops represented by identical sequences of POIs can be assigned to the same cluster of similar loops. All clusters including at least three loops are considered to identify the user’s systematic loops. The criteria defining a systematic loop are subjective and depend on the meaning assigned to the word systematic. Depending on the goal of the analysis, one may consider as systematic only loops traveled daily, or include also loops traveled once every two weeks. In this work, having used a monitoring period of only six weeks, and considering that some loops have been lost due to the limited accuracy of the tracking system, the number of three occurrences implies that a systematic loop has been traveled on average at least once every two weeks, which was considered enough to accept it as part of a user’s routines. We used a sample of ten users to test that the detected systematic loops were actually considered systematic by the users themselves.

A second loop processing step accounts for situations such as the one presented in the example below. Let us assume that user X travels twice the loop [home $\rightarrow$ holiday house $\rightarrow$ home], once the loop [home $\rightarrow$ holiday house $\rightarrow$ gas station $\rightarrow$ home] and twice the loop [home $\rightarrow$ cafeteria $\rightarrow$ holiday house $\rightarrow$ home] (see Figure 5). With the threshold set up above, none of these loops would be identified as systematic. However, we believe that the loop [home $\rightarrow$ holiday house $\rightarrow$ home] is relevant, although sometimes it has some additional stop-over, and would like to be able to identify it as a systematic loop.

![Figure 5: Similar loops treated as the same systematic loop.](image)

If a loop does not belong to a cluster of systematic loops (that is, its sequence of POIs has been observed less than three times), but it includes all POIs of another loop in the same order, the two loops are counted as if belonging to the same cluster of loops, described by the sequence of POIs of the shortest loop. Indeed, the former is considered an extension of the latter. In the example above, the sequences [home $\rightarrow$ holiday house $\rightarrow$ gas station $\rightarrow$ home] and [home $\rightarrow$ cafeteria $\rightarrow$ holiday house $\rightarrow$ home] will be counted as if belonging to the [home $\rightarrow$ holiday house $\rightarrow$ home] cluster, which will thus be identified as a systematic cluster, as it now includes five loops. This way some point of interest may be lost, yet more sequences of points of interest that are actually part of the user mobility patterns and routines can be identified. Our choice is also motivated by the fact that the activity tracker is not always capable of distinguishing between real POIs and trackpoints,
such as stations, bus stops, gas stations or other stay points where users just temporarily stop along the way to their actual destination. This creates a number of spurious loops that are, indeed, formally equivalent to systematic loops, but would not be taken into account in the assessment of the user mobility patterns. Also, in this case, processing would not be necessary (being, instead, detrimental) whenever the POIs were more reliably identified, for instance, by a more accurate GPS tracking device or by asking the user to validate them.

4.2. Identification of Alternatives

In this section, we present the algorithm developed to identify optimal, low-carbon alternatives for all observed travels.

Identification of Alternatives for Routes. In a first step, a route planner computes a set of eligible routes, by considering the different modes of transport (MOT) mentioned in Table 1. As different MOT can often be combined with a positive effect, we additionally use the OpenTripPlanner (OTP) route planner\footnote{The OpenTripPlanner (opentripplanner.org) is an open source route planner that is able to find multi-modal public transport routes.} and an implementation based on the model introduced in [8] to find multi-modal trips. The OTP uses an implementation of Round-Based Public Transit Routing\footnote{11} for schedule-based transport modes, but is able to combine this method with street-level routing to compute multi-modal journeys. As the implementation takes into account the exact departure times of public transport, we picked a representative time (the time of the route that represents the loop “best”, in the sense that it shares most POIs with other routes in the same loop and has the smallest departure time difference to them) for each systematic route when querying the OTP. The exact departure and arrival time of the resulting routes were discarded, though, and only the overall duration was kept for analysis. The approach presented in [8] is primarily used for less frequently used combinations, which cannot be handled by the OTP (such as combinations of bicycle and public transport). It essentially uses a heuristical two-step procedure, where first potential transfer nodes are computed (between different modes of transport), which are evaluated for feasibility in a second round (e.g., determining if there is actually a train connecting two potential transfer nodes). In terms of the overall framework, the concrete routing application could be replaced with any routing service. We primarily relied on OTP as it is one of the few freely available public transport routing applications (which removed any constraints on the number of alternatives we could retrieve from it). For all the alternatives retrieved by the route planning algorithms as well as for the original route, we then compute key features, such as the duration and the total CO₂ production. For each original route \( r_o \), this results in a set of alternatives \( A = \{ r_a \} \) (where \( r_a \) denotes the alternative route with its trackpoints and modes, which has an associated total duration \( d_a \) and CO₂ production \( c_{a} \)) that could be used to replace the original route \( r_o \).

In a second step, impossible and highly unlikely suggestions are removed from \( A \), if they fulfill any of the following conditions.

- Any of their modes of transport appears in the user’s list of unavailable modes of transport.
- The original and alternative routes are too similar. To assess this, we define a similarity measure between the two routes as the fraction of distances between the trackpoints of the original route and the closest trackpoints of the alternative that are below 200 meters. Two routes are considered too similar whenever the modes of the original route and the alternative coincide and the similarity measure exceeds 70%.
- The alternative does not reduce the CO₂ production by more than 5% (compared to the original).
- The total slope (in elevation) of a bicycle-based alternative is more than 1.5%.
• The duration of the alternative is too long, i.e., it exceeds the threshold $d_{th}$:

$$d_{th} = (d_o + t_{max}) - \frac{t_{max}}{1 + d_o \cdot t_{max}}$$

(1)

where the parameter $t_{max}$ represents the maximal increase in duration and $d_o$ denotes the duration in hours of the original route. In our implementation, we set the maximal prolongation of the route duration as $t_{max} = 1.2$ hours (=1:12 h). Of course, this limit is not reached for shorter durations (e.g., a route with an original length of 5 min can maximally be replaced by one of 11.5 minutes). Figure 6 shows the value of $d_{th}$ as a function of $d_o$.

![Figure 6: Threshold $d_{th}$ for the alternative duration.](image)

It would be possible to filter out routes with long distances as well, even differentiating them by MOT (e.g., preventing alternatives where one would have to walk for 10 km), but in practice such routes are already filtered out by the duration constraint, except if the original route is itself in a similar range. For this reason, we do not use any distance filter.

Finally, the remaining alternatives must be ranked according to some criteria. While more elaborated choices could be considered (especially if one wished to account for user preferences, such as wanting to lead a healthy lifestyle, or disliking some form of transport), here we simply rank the remaining routes according to the CO$_2$ produced, in order to give feedback about the sustainability potential of an alternative.

Identification of Alternatives for Systematic Loops. For each loop, we can identify alternatives in a similar way as we did for routes. There are some circumstances that have to be specially considered, though: when generating an alternative loop, the sequence of transport modes has to be respected. Moreover, the total additional duration of an alternative loop should be further bounded. Using above formula, for example, a loop of four routes of 30 minutes (a total of 2 hours) could be increased up to 3:48 hours, which is clearly unacceptable for users. Therefore, requirement 2 of Section 3.5 is satisfied by bounding the total time of an alternative loop using Formula 1 with a maximal time increase of $t_{max} = 1:24$ hours. Then, in the above loop of four routes, which originally lasted 2 hours, the total loop duration can be increased by at most 1:02 hours, instead of the original 1:48 hours.

The requirements are implemented by sequentially processing the routes of the original loop, and building a graph of options, storing, at each node, the visited location and all available modes of transport. Figure 7 shows an exemplifying graph. The blue POI depicts the user’s home, where the transport modes walk, car, bike and public transport (PT) are available. While in A) the algorithm suggests to take public transport to the first POI, in B) the use of the car is suggested. In the first case, all consecutive route segments need to be of type public transport or walk, as neither the car nor the bicycle are available anymore. In the second case, taking the car would, in principle, force the second segment to be of either walk, car or public transport type; however, since the car needs to return to the user’s home in the end, all modal choices except the car are pruned after processing the last segment.
5. Results

To get a deeper insight in the methods described above and to validate them, we have interrogated the participants of the here presented study. Having obtained a confirmation of their validity, we have then analyzed the aggregated results of the performed computations, by estimating the *overall potential for change* of the participants: such aggregated analyses can provide a (rough) upper bound of the potential effects of a successful intervention.

5.1. Survey-based User Assessment

To assess the effectiveness and accuracy of the data analyses presented above, we performed two surveys, targeting the users themselves. The first survey aimed at assessing the users’ perceptions about their overall satisfaction throughout the project, and dedicated a few questions to investigating the quality and usefulness of the individual eco-feedback reports. It was delivered to all participants, by means of an online questionnaire; between 102 and 104 responses were collected for the subset of questions regarding the eco-feedback report, equivalent to 39.5% of the sample of 261 study participants. This complete sample was recruited by means of a public communication campaign, attracting around 700 voluntary people living in two Swiss regions (the city of Zurich, a dense urban context characterized by high quality public transport provision, and the Canton Ticino, a much less dense context, characterized by urban sprawl, where car is the dominant transport mode). Filtering due to data incompleteness and for representativeness purposes led to the final 261 participants. Referring to a 7-point Likert scale (where 1 = “totally disagree” and 7 = “totally agree”), we investigated the level of agreement of users with the following sentences:

- The reports correctly identified my reference mobility patterns, in terms of percentage of use of the means of transport:
  \[ M = 5.81, \text{ SD} = 0.98 \text{ (n= 104)} \];

- The reports correctly identified my reference mobility patterns, in terms of systematic journeys:
  \[ M = 5.73, \text{ SD} = 1.26 \text{ (n= 103)} \];

- The reports suggested realistic and feasible alternatives for my systematic journeys:
  \[ M = 4.31, \text{ SD} = 1.47 \text{ (n= 102)} \];

Overall, these responses show a definitely positive assessment by users, especially about the system’s capacity of automatically identifying the means of transport and the systematic loops (for the sake of simplicity, called “systematic journeys” in the survey questions). To get a deeper insight, we developed a second survey, again administered by means of an online questionnaire, aimed at collecting the perception of accuracy by each user on each single systematic loop we had identified and on the related alternatives found. We obtained responses by 85 users, namely 32.6% of the overall sample of study participants (261). Users were shown a map for each systematic loop we had identified, presenting the sequence of points of interest defining each of them, and were asked (Q1) whether they had “traveled between the points shown in the figure, for at least three times during the first monitoring period (March-April 2016)”. Possible responses where Yes/No/Partially. Then, they were shown an updated map, also reporting route connections between such POIs, and were asked a second question (Q2): “Would you classify this route as *systematic* (namely, a route that you frequently travel)?” Possible responses ranged from 1 to 5, where 1 = “definitely no” and 5 = “definitely yes”.

![Figure 7: Example of computation for loop alternatives.](image)
The collected answers refer to a total of 651 systematic loops, by 85 different users. Question Q1 obtained responses for 550 systematic loops: 86% of them were “Yes” and 6.8% were “No”. This confirms a high accuracy in the identification of repeatedly traveled loops. Moreover, from responses to question Q2 (again obtained for 550 systematic loops), we got the confirmation that the frequent loops we had identified were generally perceived as systematic: the mean value of Q2 responses was 3.98 (standard deviation = 1.40), with 58% of the identified loops classified as definitely correct and 69% of them classified as either correct or definitely correct (Likert scores equal to 4 and 5, see Figure 8). From the comments to questions Q1 and Q2, we noticed that several users did not perceive those loops frequently traveled for shopping or leisure reasons as systematic; others have judged loops as non-systematic because they have changed their routines, but admitted that those loops where systematic during the monitoring period. This explains why the percentage of positive answers to Q2 is smaller than for Q1.

For each loop for which a CO₂-efficient alternative was automatically found, the questionnaire also showed the user the alternative (consisting of the suggested route and the means of transport between the same POIs). We identified 235 such alternatives for the 651 systematic loops (36.1%). Based on these 235 alternatives, we asked a final question (Q3): “Do you think the suggested route and means of transport are plausible?”. Again, possible responses ranged from 1 to 5, where 1 = “definitely no” and 5 = “definitely yes”. Overall, we received responses for 42 out of the 235 alternatives for systematic loops (17.9%). In this case, the perceived accuracy by the users was lower, indicated by a mean value of 3.38 (SD = 1.62) in answers to Q3. However, if we consider as plausible alternatives those assessed with a score of 4 or 5, they result in 50% of the alternatives automatically found (see Figure 8). A score of 3 (neutral assessment) is attributed to 17% of the alternatives, thus leaving a 33% of alternatives that are assessed as not plausible. Due to the way the Q3 question was framed, these answers to Q3 could at least partially be affected by a “social desirability bias” effect [34]. As a consequence, we acknowledge that the actual percentage of plausible alternatives might be slightly lower than indicated by current figures, though we are not able to quantify how much.

Figure 8: Assessment provided by survey respondents for (a) each systematic loop and (b) each alternative to a systematic loop we had identified, according to a 5-point Likert scale.

5.2. Analysis of the Overall Potential for Change

Finally, we analyzed the results obtained when applying the previously introduced methods to all the 261 users of our study. Although affected by the approximations necessarily involved in real data collection, these results give the scale of potential energy savings that can be achieved by fostering changes in the mobility behavior of comparable communities or groups of individuals. Taking into account the above limitations related to the automatic identification of alternatives, it is important to note that the results presented below should be interpreted as upper boundaries for the potential savings, which can be achieved only in the ideal situation where no other constraints limit user choices beside those applied here to the travel duration and mode availability. Further, plane trips have been removed from the collected data, as they strongly impact
Table 2: Analysis of the mobility patterns and potentials for change over all users. SL stands for “systematic loops” while NS stands for “non systematic” and refers to routes. Reported are the average values per user, if not indicated otherwise (e.g., the average user travels 387.1 km (76.9 + 310.2) per week, but could potentially reduce this to 318.8 km (45.5 + 273.3) per week).

Table 2 shows some global results obtained when applying the methods described above to the routes collected for all 261 users monitored during the project. The first part of the table gives the total number (tot) of users monitored, routes recorded and loops reconstructed, the average number (avg) of systematic loops (SL) identified for each user and the average number of times each SL is traveled (SL repetitions). One can see that on average, mobility patterns of a user can be characterized by about 5.3 systematic loops, repeated about 5.2 times in a six week period. The second part of the table presents the distances traveled, the energy consumed and the CO\textsubscript{2} emitted each week, reported as the mean (over all users) of the average weekly values. All values are separately reported for systematic loops (SL) and for all the other non-systematic routes (NS). The “alternative” columns compare these values to what could be achieved by using the alternative solutions identified by our algorithms. Finally, in its third part, Table 2 analyzes the overall modal choice, i.e., the fraction of the total distance (summed over all users) traveled with each mode of transport. Figure 9 provides more detail about the distribution of such indicators among users, showing box-plots for the modal choices and the average distances traveled, energy consumed and CO\textsubscript{2} emitted weekly. Each row shows both the originally tracked data and the computed alternatives, considering the systematic loops as well as all remaining non-systematic routes. The individual dots in the figure are outlier users, while the boxes represent 50\% of all users. Looking at the first rows, for example, one can see that the parts of systematic loops that are traveled by car could be replaced by more sustainable means of transport in many cases, while the non-systematic car routes have a much smaller potential for replacement (often due to the fact that they cover long distances to places far away from public transport stops, preventing both the use of a bicycle as well as public transport). It can be seen how a large part of the potential energy savings stem from a transition from car to public transport, bicycles, and walking. Energy savings, however, can also come from a decrease in the weekly traveled distance. These savings come from situations similar to the example in Figure 2 where the detours of a public transport journey are replaced by a more direct and thus shorter bicycle path.
6. Discussion

As presented in the previous section, the user survey highlights that the systematic loops found were generally of satisfactory quality. It is more difficult to propose meaningful alternatives to users, partially shown by the fact that we “only” found 235 alternatives for 651 systematic loops, partially by the lower score of Q3. There are multiple reasons why it is impossible to find suitable alternatives for all systematic loops, but we would like to highlight two important ones here. First, as the loop length increases, it becomes more difficult to find a suitable alternative, because each segment of the loop requires a suitable...
alternative, and in sequence they must satisfy the mode availability constraints. This implies that finding
alternatives for individual routes is much easier than for complete loops. Second, in several surveys and
interviews we conducted, we found that our study sample is biased towards people who regard ecological
sustainability as a requirement for the future of humanity (pro-environmental attitude). Such people seem
to be naturally inclined to participate in experiments that try to make behavior more ecologically sustainable.
This attitude manifests in their systematically traveled loops, which are shorter and often covered more with
sustainable modes of transport than their non-systematic counterparts. In our two study regions, we found
that non-systematic routes are 3.4 (Zurich; urban) and 4.1 (Ticino; rural) times longer than routes appearing
in systematic loops. In Ticino, 74% of the routes of systematic loops are traveled by car, compared to 71% of
the non-systematic ones. In Zurich, however, 36% of the routes of systematic loops are covered by car, in
contrast to 51% of the non-systematic ones. This means that in general we have higher chances to find a
promising alternative for a non-systematic route.

To explain the lower score for the automatic identification of alternatives (Q3), one should consider
that our methods do not include many factors which may influence mobility choices and that the analyses
performed are necessarily affected by the inaccuracies of the tracking system. For instance, if a visited place
is not recognized properly by our method, the suggested alternative does not visit this place of actual interest
for the user. This in turn will cause the user to assess it as “not useful”. Moreover, users might have specific
constraints, related to the need for accompanying family members, for having to carry weights, or for being
capable of suddenly changing schedules and routes due to family or job reasons. Such specific elements are
highly user dependent and cannot be taken into account by the system. Therefore, even if theoretically
speaking a more efficient alternative from the point of view of CO\textsubscript{2} emissions is available to connect the
visited points of interest, specific needs of single users might lead to assess it as unfavorable and not realistic.
As this is highly dependent on the context of a user, it can also happen that someone classifies an alternative
as feasible, but still is not able to use the alternative in a majority of cases. While in the here presented
study we only looked at the overall CO\textsubscript{2} and energy savings, it would be interesting to analyze the individual
route choices of people over a certain time span. In a “living lab” without particular constraints on the
participants, observing such changes is difficult, though, as people change their places of living and work,
their personal circumstances, etc.

The eco-feedback provided only marginally accounts for individual differences between users. As mentioned
above, personal needs and preferences cannot always be recognized from automatic tracked data. As such,
some reasonable requirements constraining the modal choice have been assumed to apply indiscriminately to
all users. A more involved interaction with the users (e.g., asking people to state the purpose of each trip), as
well as improved tracking accuracy, would improve the accuracy of the suggestions and allow for some degree
of personalization of the feedback. In retrospect, a more accurate tracking app, allowing the inclusion
of travel survey techniques (e.g., asking for additional data about each trip) could have been beneficial,
in particular for the derivation of more meaningful alternatives. Relying on Moves not only discouraged
asking users for additional travel data, but also meant relying on the given accuracy for the identification
of points of interest, and thus recognizing of loops. Ultimately, these circumstances also influenced the upper
bounds presented in Table 2 and Figure 9 as the tracking inaccuracies can lead to points of interest not
being recognized, and thus not being respected in the computation of alternatives.

However, increasing the frequency of interactions between the mobility tracking app and its users would
need a careful assessment. In fact, in order to make smartphone-based tracking a viable solution for real world,
large-scale mobility monitoring, very little or no interaction with the user should be requested. This also
made it impossible to employ a “traditional” travel survey (where people manually enter traveled journeys
into a system). Additionally, in projects and studies aimed at assessing the effectiveness of behavior change
interventions in controlled trials, namely by comparing mobility patterns before and after an intervention, and
with respect to a control group, any interactions between the app and its users would potentially influence
their behavior and, thus, the outcome of the study. As shown by high drop-out rate registered through
the three phases of our project and by final interviews we performed with project participants, there is a
tangible risk that users feel under pressure by the interaction effort requested by the app and, to avoid it,
directly opt for quitting app use. Therefore, the availability of a smartphone-based tracking tool requiring
none or very limited interaction with the users would be largely beneficial also for other research purposes,
besides pure mobility tracking. This induces that the tracking system essentially includes a reliable transport mode classifier, as this would remove the need of asking for the actual transport mode used (due to the unavailability of a highly accurate classifier, in our study, users were asked to manually validate each recorded route).

The eco-feedback presented here was primarily delivered to users in terms of a static report as well as with the use of several gamification elements. Especially the potential for change can be used as an individual starting point for goal setting, concrete suggestions, or rankings of users, as these elements are highly user- and context-specific (e.g., creating a leaderboard of “most eco-friendly” users would be heavily biased towards people living in cities or having short journeys to work). Alternative or supplementary choices to the above presented eco-feedback include monetary and educational (any impersonal information about the ecological sustainability of certain travel behaviors) incentives. Within this study we focused on automatically delivered personalized feedback for various reasons: monetary incentives are purely extrinsic motivators, and are thus known to only show temporary effects [4, 15]. Additionally, while educational incentives are especially valuable for uninformed people, they show smaller effects for people who are already aware of a certain problem, and are considering a change, though are still caught in a typical “attitude-behaviour” or “value-action” gap [26, 1]. As several surveys and interviews accompanying the study showed, our participants were already inclined to optimize their travel behavior with respect to ecological sustainability, but did not know how to improve it or lacked discipline in changing their behavior.

7. Conclusion and Outlooks

We have presented a method to analyze automatically tracked location trajectories and provide eco-feedback to their users. The method builds upon the identification of systematic mobility patterns in user trajectories and the suggestion of alternative modal choices for such frequent travels. In particular we have focused on the identification of systematic loops starting and ending at a user’s home. We argue that these loops build a solid foundation to compute eco-feedback, as they reflect the actual mobility needs of each user, allow providing reasonable and feasible solutions, and offer a good level of abstraction to communicate to users. For each loop, the proposed approach computes an alternative optimal solution that minimizes the CO$_2$ emissions while respecting a set of constraints (e.g., on the total duration of the travel) that can make it acceptable to the average user. This approach can effectively be applied in other countries than Switzerland, where we originally developed it, provided that country-specific CO$_2$ emissions factors for each transport mode are considered, for example by retrieving them from the ecoinvent Life Cycle Assessment database [52].

The generated eco-feedback is mostly of educational nature, i.e., it does not create an incentive for a person to change her mobility behavior, but simply shows possible ecological gains from such a change. Of course, this assessment can be used as a base to create incentives (e.g., by building gamification elements on top of it), thus making the suggestions provided to users more effective [49]. Alternatively, behavior change could be fostered by pointing out benefits of modal changes. This can include time for other activities (such as reading or working on a train) as well as economic aspects, ranging from simply showing the user how much money she would save by performing a certain behavior to actively encouraging certain behavior by travel vouchers, benefits, etc. In order to give people more meaningful travel alternatives and thus better eco-feedback, however, more accurate tracking technology as well as more detailed data about the individual journeys (e.g., the purpose, or additional constraints like luggage or people the user is traveling with) is required. A challenge for future research activities is therefore to look for a closer collaboration with app users, in order to better understand their mobility needs and receive a more detailed feedback about the quality of loop recognition and the applicability of alternatives, while at the same time trying to limit the amount of compulsory interactions with the app as much as possible.

Another line of research can concern the method itself. Transport options such as carpooling or car- or bikesharing make viable alternatives for many mobility demands, which were not included in this work. As the exact journeys people travel are known, an automated assessment of possible carpools and shared vehicles would be possible. This could be used to create eco-feedback in the form of suggestions to work together with other app users. Another extension could consider rescheduling user’s activities, in order to create an optimal daily schedule in terms of mobility and sustainability. Again, this is highly dependent
on various individual factors (e.g., trip purpose, or interaction with other people’s schedules) which would require a much stronger interaction of the user with the app.

Concerning the eco-feedback itself, its immediacy plays a big role in making it an effective tool for fostering more sustainable mobility behaviors [13]. While in our case the feedback was delivered passively after a certain period of time, the prompt recognition of a starting journey and the real-time suggestion of an alternative could greatly increase its effectiveness. The main challenge in this real-time approach is to anticipate the intentions (e.g., the destination) from the tracked trajectories of a user. This research area has recently received a lot of attention though, and will increase in importance as location based services get more and more prevalent.

References


[32] Harvey J Miller and Jiawei Han. *Geographic data mining and knowledge discovery*. CRC Press, 2009.


[52] Simon Wells, Paula Forbes, Judith Masthoff, Silvia Gabrielli, and Antti Jylha. SUPERHUB: integrating digital behaviour


