Aspects of Personal Navigation with Collaborative User Feedback

Harald Holone
Østfold University College
Halden, Norway
harald.holone@hiof.no

Gunnar Misund
Østfold University College
Halden, Norway
gunnar.misund@hiof.no

Håkon Tolsby
Østfold University College
Halden, Norway
hakon.tolsby@hiof.no

Steinar Kristoffersen
Østfold University College
Halden, Norway
steinar.kristoffersen@hiof.no

ABSTRACT

Inspired by systems based on user generated content, we have developed a prototype named *OurWay*, a collaborative route planning system utilizing user feedback (rating of route segments) to provide quality routes adapted to the users' abilities and needs. We report from an indoor experiment where users in wheelchairs solved navigational tasks with our prototype. Log data, observations, and interviews serve as a basis for discussing the feasibility of the OurWay concept. We find that OurWay yields better routes for all users with aggregated route segment ratings produced throughout the experiment. However, ratings were largely produced by each individual to accomplish a selfish goal, namely that of solving a navigational task. In this respect, rating can be seen as a *by-product* of use, rather than as an intentional action on behalf of a community.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Personal navigation, Collaboration, Accessibility, Routing

1. INTRODUCTION

In this paper we present findings from the second stage in an ongoing project, OurWay, where the main goal is to explore navigation and wayfinding based on sharing the users' experiences. The OurWay concept is comprised of mobile, map based clients and a central server. The users are encouraged to evaluate and report the accessibility of locations along routes suggested by the system. Based on the aggregation of user ratings, the central server calculates the best route between two locations.

NordiCHI 2008: Using Bridges, 18-22 October, Lund, Sweden.

Information technology is fast, cheap, and precise enough in terms of positioning to have become a valuable navigation aid. GPS navigators for motorists have become ubiquitous, and it is natural to extend the idea to pedestrians and other non-vehicle users. There is, however, a lack of data to support pedestrian navigation, and this becomes even more evident for users with mobility problems.

Some routes, even through buildings, will not be traversable for a wheelchair user, e.g., if the proposed route is using the stairs. A narrow elevator may fit a smaller wheelchair, but not a larger, electric one. Carpets, thresholds and doors are easily traversed for an able-bodied person, but may be uncomfortable or render the route unusable for others.

As a potential solution to this challenge, we have proposed a collaborative route planner. With OurWay, the users rate segments of suggested routes in a collaborative fashion, to provide traversable routes adapted to a groups needs and preferences.

In the first stage of the project, we performed a proof-of-concept evaluation. A combination of field work and desktop simulations produced indications of the potential of the application [12]. In particular, we demonstrated that a relatively small number of ratings from users yielded good routes in an urban environment. Good routes, in the eyes of the users, typically avoid obstacles, and favors locations with positive ratings.

Our initial research focused on the algorithmic and technological aspects of the system, and thus left us with questions about how and why people would rate locations along the suggested routes. Hence, we decided not to target UI issues, but rather focused on the feasibility of the concept. More specifically, we wanted to shed light on the following aspects of collaborative rating of accessibility:

Effectiveness: Given a small group of users in a close-to real life setting, will their aggregated segment ratings yield satisfactory routes?

Efficiency: How does the number of segment ratings, and the efforts involved in producing them, affect route quality?

Satisfaction: Do the users perceive the service as useful, and how do they rate route quality?

A small experiment like this has obvious limitations in terms of generalizability. However, we are confident that answering these questions at a micro level will yield significant guidelines for further development and exploration of the OurWay concept.

In the next section, we give a short outline of related projects. The OurWay prototype is described in Section 3, while research methods and experiment design are treated in Section 4. Our findings are presented in Section 5 and discussed in Section 6. We close our paper with some conclusions and directions for future work.

2. RELATED WORK

In our series of studies of the novel OurWay concept, we draw inspiration and knowledge from several fields of research and development. Early commercial efforts in pedestrian navigation include the pioneering DoCo-Navi [28] and the later KDDI's EZ Navi Walk [17]. However, commercial solutions are not frequently treated in the research literature. The major bulk of academic work touching on pedestrian wayfinding seems to be related to the domain of mobile guides (see for instance the survey EU based projects in [13]).

Pedestrians are a highly diversified group, with respect to abilities and preferences. Hence, the need for *personalized* route planning, i.e., that the route planner adapts to the user's specific needs and desires, becomes obvious. An example of personalized car navigation is found in [3]. Kawabata et al., leverage context dependent data to generate optimal pedestrian routes according to the user preferences [16]. Wuersch and Caduff remind us that pedestrians are not confined to a fixed network of streets and sidewalks, and treat routes as a sequence of waypoints [29].

In particular, when catering for people with mobility problems, the heterogeneity of the user group becomes a central challenge when trying to devise navigational aids. As an example of the scarce supply of work on this topic, Karimanzira et al., investigate the use of machine learning techniques to generate routes customized for physically disabled pedestrians [15].

As research on collaboration and recommender systems has matured (such as [11]), little attention has been paid to take advantage of these mechanisms in pedestrian navigation. An exception is McGinty and Smyth [21], using multiple agents sharing experiences to create a distributed case based reasoning system. Another proposed approach is to collaborate by sharing clues, either through direct participation [6], or indirectly through image annotation on wiki-style maps [5].

A main prerequisite for computing personalized routes is knowledge of the user's preferences. Haigh et al., propose that the user can rate routes in order to decide whether to reuse existing routes or find alternatives in unexplored territory [10]. Akasaka and Onisawa let the users classify roads according to a detailed schema, use fuzzy measures to derive the users' preferences, and assign attributes to roads based on detailed user input [1, 2]. Examples of explicit pedestrian modeling are found in [25]. On the other hand, Rogers and Langley claim that explicit user modeling may be too resource demanding, and will give too few assurances of accuracy to be worthwhile [24].

In the *MAGUS* project, a comprehensive level-of-service (LOS[1]) model for wheelchair users is developed, based on questionnaires, interviews, observations and physical measurements of starting and rolling resistance [4]. The resulting system is a desktop application, aiming to assist new users and enable better navigation for existing users, and as a means for planners. However, Sobek and Miller point out that the detailed LOS model would be extremely costly to establish and maintain, and that the application requires too much time from the users [27]. Based on these observations, Sobek and Miller present an alternative system for route planning for disabled pedestrians, called *U-Access*. They propose simplified models of both level-of-service and users, claiming that this still generate good results.

In OurWay the approach is even simpler. First, we allow the users to organize themselves based on self-identification, creating groups we can assume share abilities and preferences. Second, we let the users collaboratively generate a simple LOS model based on ratings of edges in the transportation graph.

Personalized routes, however for bicyclists, is also the focus of a project from Priedhorsky et al., [23], who rely on the wiki approach to create a personal route-finding tool for cyclists. Based on interviews, relating to a designed, but not yet available service, users were significantly positive on the issue of *contribution*, which is crucial for the sustainability of collaborative systems.

Much of the research cited so far share a focus on implementation and proof of concept, and the usability focus is, if present, marginal. However, a significant strand of mobile HCI research is concerned with design, implementation, and evaluation of mobile guides (see the five international workshops on HCI in mobile guides, e.g., [8]). Some of the applications in this category offer pedestrian navigation and wayfinding functionality. However, the HCI aspects addressed in this field of research are mainly concerned with traditional device interaction issues. In contrast, our main focus is on the service level feasibility of the concept as such.

Last, but not least, we would like to emphasize the role of available geospatial data in general, and road networks in particular. The cost and complexity of acquiring necessary underlying data for a navigational service are effectively barring creative and efficient applications development outside the premises of the established commercial and governmental actors. Our project would not have been feasible without the tools and services from the OpenStreetMap project[2]. The impact of *volunteered geographic information* is elaborated on by Goodchild in [9].

3. THE OURWAY PROTOTYPE

The OurWay prototype consists of a route planning server and a client application running on a mobile phone. Additionally, a web based visualizing tool allows for retrospective walk-through of the route planning activities, such as segment ratings and re-planning of alternative routes.

The server employs a standard A* algorithm to find the lowest cost path between two nodes in a geographical network. Feedback provided by users applies weight to edges (route segments) in the network, thus providing the adaptability in the system. In the current implementation, when more than one rating exists for a segment, the most recent

[1]The level-of-service concept is frequently used in vehicle transportation research, and covers aspects such as suitability and efficiency.

[2]http://www.openstreetmap.org

one is used for the route calculation.

Multiple (potentially disagreeing) ratings for a segment and the temporal aspects of some obstacles can be handled in a number of ways. Our choice of a wiki-style approach in the current implementation is rather arbitrary, and can easily be modified. Time-based fading and average rating values are two viable alternatives, and there is also potential benefit in visualizing other user's opinions in the map.

We have deliberately made the granularity of feedback quite coarse; only allowing for three different ratings. Users can choose among *good*, *uncomfortable* and *inaccessible* when they decide to provide feedback to the system in the form of route segment ratings. When taught how to use the system, the users were explained that *good* could typically be applied to a segment (part of a route) that stands out as better than neutral. *Uncomfortable* was to be regarded as an obstacle that is passable, but can be troublesome (say, a door opening the wrong way), whereas *inaccessible* is an obstacle which is not passable, such as a flight of stairs, keeping in mind that our users in this experiment were situated in wheelchairs.

Previous *outdoor* experiments led us to choose these three levels as opposed to, say, five levels, mostly because there is too much individual variance to extract any meaningful information from the extra two levels, and also because we wanted to make the interface as simple as possible for the user. We also suspect that keeping the number of choices to a minimum potentially lowers the threshold for bothering to rate a segment of a route.

Our previous OurWay experiments have been located outdoors, which allowed for the use of GPS to position the user. However, as this experiment was taking place indoors, GPS positioning is not a viable choice. Neither was any other indoor positioning system installed in the building, which required us to implement manual positioning by the users themselves. The user would look at the map on the screen, and see her own position marked with a red cross hair. As we moved around in the building, the user would "push" the cross hair with the mobile phone joystick, much like pointing on a paper map with a pencil. In a similar fashion, the user was required to select the current floor.

Today, the geographical networks which constitute maps are not readily available. Google Maps is a well known example, and most countries have local map providers and guides, which are available via the Web. These services are not easily used as a point of departure for location based or context-dependent mobile annotation applications, though, partly because they generally only provide map images, and not access to the underlying geographical network, and partly because of the restrictive licenses under which the map images are made available. One particularly pertaining exception is the OpenStreetMap (OSM) effort.

OSM offers an open standard data format for describing a geographical network, and there are suitable tools available to support the data creation and maintenance from a desktop platform. The resulting network can easily can be augmented with data from other sources; one example is how we apply feedback from OurWay users to the geographical network of edges and nodes. The geographic information from OSM is made available under a Creative Commons license. The OpenStreetMap tools and infrastructure serves as a starting point for the application we are describing in this paper.

As might be expected, no geographical network existed

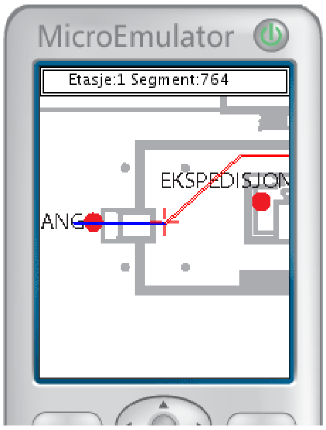


Figure 1: Client screen shot

for the interior of our campus building. We chose to use the OpenStreetMap tools for creating the network, and then imported the network into the OurWay server. On the OurWay client, a set of map tiles for each floor of the building was installed as part of the application, requiring no fetching of map images during use.

The client (see Figure 1) is capable of requesting from the server a route from one node to another, and the calculated route is returned to the client and rendered on top of the building map. The segment (edge of the transportation graph) of the route closest to the cross hair is highlighted, and it is this segment which is rated when the user provides feedback to the system. For the purpose of this experiment, six predefined tasks (comprised of start and end nodes) are available from a menu in the client.

The mobile client has a basic set of functionality required to cater for the specific research objectives of this work. Hence, a thorough evaluation of algorithmic issues and the user interface is outside the scope of this paper.

4. RESEARCH FRAMEWORK

4.1 Method

At this stage in the OurWay project, we are primarily interested in shedding light on the feasibility aspects of the concept. Thus, we are not particularly concerned with device interaction usability, but rather with the value of the system in a close to real world setting, as perceived by real users [20]. To achieve our goal, we employ an established usability framework as a tool for discussing our findings. We do, however, keep in mind not to expand the construct beyond reasonable limits [22].

Practices for conducting usability assessment of mobile applications is a frequently debated research issue. A recurring theme of discussion is field tests vs. laboratory experiments. In a comprehensive survey of mobile HCI methods, Kjeldskov et. al., find a clear bias towards implementation oriented approaches. Furthermore, in the cases where evaluation was on the agenda, the preferred setting was the laboratory. They argue that this might inhibit future development of the HCI field [18]. On the other hand, several papers report that field tests, compared to laboratory experiments, yield minimal added value, see e.g., [19].

In our case, the initial proof-of-concept evaluation was carried out in a controlled field and laboratory setting. Thus, to gain further insight, there was a strong call for targeting users in a setting as close to real life as possible.

We finally settled on the case of indoor campus navigation for users in wheelchairs. In this setting, we were able to create a realistic, dense and constructed environment, in particular enabling us to study a significant number of opportunities for and instances of route segment ratings in a short time span. We also decided on a traditional suite of tools for data acquisition: data logging, observations on-the-go, and semi-structured user debriefings.

As a vehicle for designing the experiment and structuring the following discussion, we took advantage of the widely adopted ISO definition of usability [14]:

> The extent to which a *product* can be used by *specified users* to achieve specified *goals* with *effectiveness*, *efficiency* and *satisfaction* in a specified *context* of use.

We specify the usability components, including criteria and assessment tools, as follows:

Product: The overall OurWay concept, with emphasis on the high-level functionality offered by the available client-server based implementation.

Users: Our participants were confined, by us, to self-propelled wheelchairs, thereby taking the roles of temporarily disabled people.

Goals: Successful and comfortable navigation between two given locations.

Effectiveness: In our case, to answer the fundamental question "Does it work?", we needed to investigate if the aggregated annotations of a small user group yielded satisfactory routes. To assess this, we compared data logs of segment ratings from the users and statements from debriefings.

Efficiency: Our main efficiency measure had two components: 1) the number of location ratings relative to experienced quality of the routes, and 2) the experienced cost of producing accessibility ratings. Again, data logs and debriefings were used in the analysis.

Satisfaction: Two main indicators were selected: 1) perceived route quality, and 2) perceived overall usefulness of the concept. We relied on debriefing statements to shed light on user satisfaction.

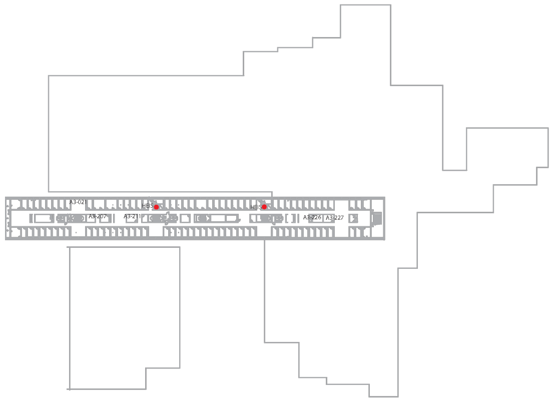
Context: Self transportation between indoor locations, in a normal working day situation.

A detailed description of the experiment is given in the next section.

4.2 Experiment

The route planning experiments took place indoors, in our 30,000 m^2 campus building in Halden, Norway. The building has five floors, and the geographical network for the building is regarded as complete. The floor layout is quite diverse, as can be seen in Figure 2. The experiment tasks were selected to span the building from basement to top floor.

(a) Ground floor



(b) Second floor

Figure 2: Floor plans of the Remmen campus building

The participating users were given a wheelchair, and presented with a realistic scenario involving wheelchair use. An accident which temporarily confines one to a wheelchair is something anyone could experience, and this was the context the participants were presented with. More importantly, the light weight wheelchair and (after all) normal capacity of the users made it possible for them to focus on the activities related to negotiating obstacles and providing ratings of accessibility along the route.

The first participant started with a neutral network, with no existing segment ratings. Thus, the first routes planned only considered the shortest distance to the target, and not the accessibility of the route segments. Later participants enjoyed the benefit of ratings made by others during the trials.

We chose six representative assignments which all the participants were to complete. As an example, one of the tasks was to get from the reception area to the student administration offices.

Nine participants (two male, seven female, age from 29 to 60) took part in the experiment, with an average of about two hours spent per person. None of the participants had prior experience as wheelchair users. They were presented

Figure 3: User rating an inaccessible flight of stairs

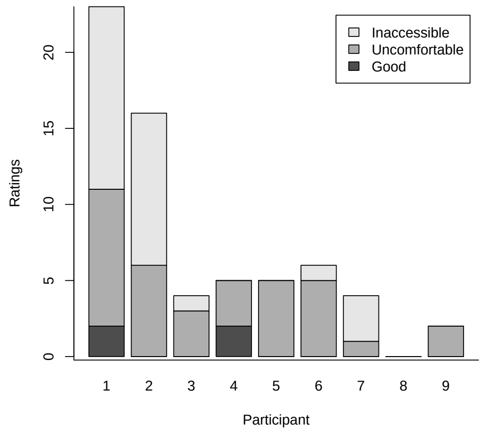


Figure 4: Segment ratings over time

with the scenario that this was their first day in the building in a wheelchair, suffering from broken legs after a car accident.

Each participant was given a short tutorial on how to use the equipment before we started. They were explained the three levels of feedback available for them to use (*good*, *uncomfortable* and *inaccessible*), and were encouraged to rate segments of the route when they saw fit (see Figure 3). The OurWay concept was explained, and they were told how their feedback would influence the routes produced by the system.

The six tasks constituted a round-trip of the building, that is, task B started where task A ended, and task F took us back to the start of task A. The start task for each participant was chosen in a round-robin fashion. Each task was presented for the user with a description of purpose and destination.

After each task, a short debriefing took place. The participants were asked to reflect on *route quality*, *usefulness of service* and *product quality*, and asked how they would have solved the task without technology.

The experiment coordinator took the role of assistant, both for pushing the wheelchair and to assist with the mobile technology when needed. The participants were encouraged to give potential obstacles a try for themselves before requiring help from the assistant. To complete the route planning team, two observers made observations and secured photos and video.

All accessibility ratings and route planning activity was logged by the route planning server. This allowed us to retrospectively inspect the routes and the feedback provided by the users.

We are of course aware of the limitation of data occurring "naturally" from an experiment such as ours, but as we shall see, the artificial and directed setting of the fieldwork contributes to strengthen the findings. The "purpose" of the subjects involvement was to study the use of the application, as a proof of concept and especially concerning the usefulness of the collaborative route planner. The subjects were taught how to provide feedback to the system immediately before their session, and so the aspect of functionality concerning the input of ratings was very fresh. In other words, we tried to establish a temporal community of practice, but as we shall see, this was soon neglected by the users.

A while after the experiments took place, we had a retrospective walk through of the experiment with one of the participants. We have used reflections from this user to shed light on some of our observations during the experiments.

5. FINDINGS

We first present an overview of the accessibility ratings provided by the users and some of our observations related to rating practice. Then we present results from the debriefing sessions after each task and reflections from the users about the system's usefulness.

5.1 Route segment ratings

We have, in previous work, established route convergence as the point at which no segments of a given route does not encourage further feedback. That is, there are no more absolute obstacles to be avoided, and no alternatives are available for avoiding the uncomfortable segments of the route [12].

Figure 4 shows the ratings created by the users. A total of 65 ratings for 42 different route segments were produced during the experiments; four *good*, 34 *uncomfortable* and 27 *inaccessible*. The first two participants did most of the hard work, encountering 12 and 10 inaccessible segments respectively.

Participant 3 met only one inaccessible segment, and the fourth and fifth participants went through the tasks A-F without experiencing one single inaccessible route segment.

Participant 6, familiar with the building, decided during one task to rate a segment taking him through a door leading into a hallway as uncomfortable, because he didn't *want* to go that way. By knowing the building and understanding how the route planning system works, he was able to adapt the route to his liking, in a way that "conflicts" with the intended rating practice.

Due to a system bug, participant 7 was given a route which traversed a flight of stairs, leading to the three segments rated as *inaccessible*.

Although ratings of uncomfortable segments seems rather unpredictable, there are some obstacles that received a fair amount of "votes" from the users. In particular, segments leading through two specific doors each were rated as uncomfortable seven times; both doors are rather heavy and especially difficult to traverse in the direction where the door opens away from the user.

It is also worth noting that rating a segment as an absolute obstacle and subsequently requesting an alternative route gives an immediate reward (a new, better route), whilst rating an inconvenient segment requires a certain degree of altruism, since there is no direct benefit for the user. After passing an uncomfortable segment, there is no need to request a new route. Even if a new route is requested, the newly added rating would not influence the route from the current location to the target since the obstacle has already been passed by the user.

5.2 Observations

The two first participants experienced more obstacles than the subsequent users, since the geographic network didn't contain any accessibility ratings to begin with. On average, participant 1 and 2 provided between three and four segment ratings for every new route they were asked to follow through the building. Although they knew they were participating in an experiment, they got quite irritated with the system when the same *type* of obstacle, typically an inaccessible staircase, appeared over and over again. When participant 1 on his way from the basement to the ground floor was sent to a staircase instead of the lift, he was frustrated. This was the fourth time that the system suggested a route where the participant ran into an inaccessible staircase. In the previous task he had rolled the same route in the opposite direction, and he had used the lift to get to the basement. Therefore he could not understand why the system had suggested another and impossible route this time and he complained:

> "The system knew I was going to the ground floor. We used the lift to get here. Despite that, the system wanted me to use the stairs this time."

A bit resigned, he rated the staircase segment as inaccessible. This was unavoidable because it was the only way to get an alternative route, which this time took him to the lift.

When participant 3 completed his tasks, he had only run into one inaccessible obstacle. All absolute hindrances relevant to our navigational tasks were now identified. The participants were enforced to rate inaccessible obstacles by using the system (and complying to the experimental context), however the system did not encourage them to rate what they experienced as uncomfortable or good route segments. The only motivation they were given was an explanation in the introduction to the project and the OurWay concept, when they were told that their segment ratings would be valuable for other users. What we observed was that the participants were rating uncomfortable and good segments occasionally and without any rational purpose.

For instance, participant 5 stopped to rate a segment taking her across a bridge between two sections of the same floor, explaining to us that she felt it was narrow, although she quite easily passed it in the wheelchair. The same user also rated a segment through a door as uncomfortable, arguing that the door was heavy to open (keeping in mind that she had to do it from sitting down in a wheelchair without mechanical brakes), but only after it had been emphasized by one of the assistants that "Now, you can try this on your own." When returning by the next route through the same door, the segment was rated as uncomfortable once more.

Working her way through the third route, however, participant 5 finds the way easily and displays excellent spatial understanding of a building that she claims not to know very well. The way-finding application works well. Route segment rating in order to enrich the underlying network for other users seems, however, to have slipped out of scope. She struggles for a while to get through another door, which is a fireproof closure between sections and thus very heavy. This is the biggest obstacle so far, yet she does not stop to rate the corresponding route segment. The door in question is one of the two doors which was identified as uncomfortable on seven occasions by different users. Coming back the same way, on yet another route in the experiment, the door is still awkward, however the user provides no feedback to the system. Traversing a small incline in the corridor, which she quite easily surmounted, she remembers that the application accepts route segment ratings. Whilst rating the segment corresponding to the graded corridor she also retraces her path in the map and rates the heavy door segment as uncomfortable.

The participants were reluctant to provide ratings for any other purpose than inducing the system to provide an alternative route (and many more examples manifested themselves in the relatively small number of subjects that took part in our experiment). They were focused on solving the navigation task, getting from A to B, and not on sharing experiences with other users of the system. When participant 1 had fought his way trough several unmanageable doors he stated:

> "I will not comment on this."

He was neither obliged, nor motivated, to rate the segment leading through the uncomfortable door.

5.3 Debriefing

After each completed task, the users were interviewed about the route they had followed, and how they appreciated the quality of the route, the usefulness of the service and the quality of the product. They rated each of these aspects on a scale from 1 to 5 (5 being best), and were asked to elaborate on their ratings.

5.3.1 Route quality

Figure 5 shows how the users reported the conceived route quality over time. Surprisingly, the rating is relatively high, even for the first two participants.

Often, the quality of the route is regarded as good, even though the users in retrospect remembers inconveniences along the route. The users justify this by assuming that the route planner "knows best", and that they have been guided along the shortest path to solve the task in question. Participant 6 gave a typical explanation:

> "I think this must have been the quickest route for solving this task. There was a narrow elevator, and the white door in the basement was heavy. I'll give it 4 for quality."

The first participant in general showed a great deal of enthusiasm for the concept, and this is reflected in his quality

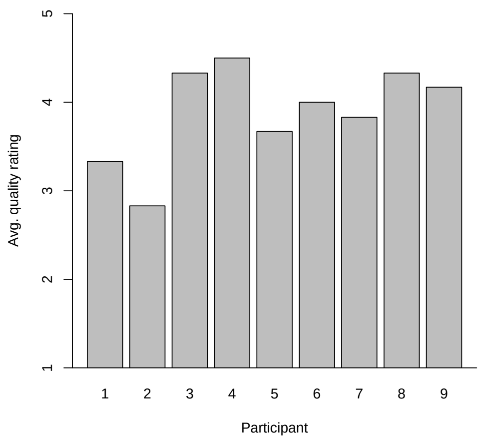


Figure 5: Route quality over time

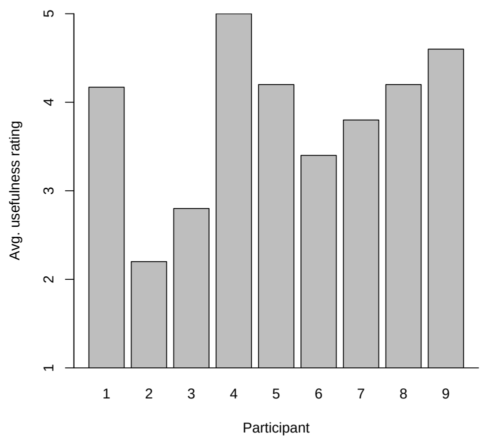


Figure 6: Usefulness over time

ratings. Participant 2 was from the get-go rather engaged from a wheelchair users perspective (she has at least one family member who is a wheelchair user), and this shows in her ratings. For example, after one of the tasks she asks:

> “Is this a doorsill finder?”

The comment was given tongue-in-cheek, but it shows her awareness of uncomfortable obstacles, perhaps also on behalf of other wheelchair users.

Towards the end of the series of experiments, the comments are generally more positive. The following quote serves as an example:

> “This was an efficient route. No obstacles encountered. Excellent!”

Still, the users often mentioned uncomfortable doors and narrow elevators without letting this affect their rating to any large degree.

5.3.2 *Usefulness of service*

The trend of the usefulness ratings is positive, as Figure. 6 shows. Even when compensating for the enthusiastic first participator, the conceived usefulness correlates well with the number of absolute obstacles the users meet during the task routes.

The perceived usefulness is also, not surprisingly, linked to how well the user knows the building. For instance:

> “I could have done this myself, I know my way around the building.”

Which leads the user to rate the usefulness to 2 for one of the tasks.

Some users rated the usefulness of the system high on some task because the route planner would lead them directly to a specific office, thus removing the need to ask strangers for directions. Removal of uncertainty was also appreciated by others:

> “This was very useful, since the signposting [in the building] is both sparse and in some cases plain wrong. It also takes me to the correct floor.”

When thinking about usefulness, some users were immediately taking the perspective of people not familiar with the building. In some cases this lead them to rate usefulness high, even though they say they could easily have managed without the route planner since they are familiar with the building.

On the other hand, the two first participants had to deal with a network with no prior accessibility ratings, leading to troublesome routes. Participant 1 reflected on the usefulness of the service. He argued that it was unacceptable for a user to be given inaccessible routes over and over again. He says:

> “If it hadn’t been an experiment, I had just asked someone to help me.”

Participant 2 expressed the same frustration after fulfilling a troublesome route in the basement where she ran into several obstacles. She was asked whether or not she could have done it better without the technology.

> “I would always mange to scream in order to get someone to help me and push me over the thresholds. And I could call the porter and my mother.”

Participant 1 further explained why he did not always stop to rate uncomfortable route segments:

> “[This is because] when I meet the same kind of obstacle again and again [...] The seventh time I run into a threshold which is too high to pass, I get fed up.”

6. DISCUSSION

The OurWay collaborative route planner is a prototype in development. Our previous work indicates that the core idea of a collaborative route planner is promising, and that our simplistic approach with three levels of user feedback is sufficient for rapid route convergence. What we set out to do in this piece of research was to investigate the feasibility of the *concept*, based on observations of a group of users in an experimental setting.

By examining log data from the route planning exercises, augmented with observations and interviews with the users, we have established an empirical foundation for evaluating the OurWay concept. Looking now at the lessons that can be learned from our experiment, we structure the discussion according to the analytic framework of effectiveness, efficiency, and satisfaction, which was presented in Section 4.1.

6.1 Effectiveness

With previous clear indications that the concept could work if "everybody play by the rules", we wanted to see if it could also apply when the users were not intimately familiar with the inner workings of the system:

> Given a small group of users in a close-to real life setting, will their aggregated segment ratings yield satisfactory routes?

By satisfactory routes, we refer to our definition of route convergence, i.e., routes that do not encourage more user feedback or has no more room for improvement in terms of avoidable obstacles.

The answer is a clear *yes, it does work*. After only two participants had completed their tasks, the route suggestions produced by the system were practically free of absolute obstacles. Given the size of the geographical network, this was the expected result. In the longer run, and in an outside scenario, however, we would not expect this to be much different. With the practices of segment rating remaining the same, there is no á priori difference between an indoor and an outdoor scenario. More interestingly, the route convergence happens *despite* the users providing feedback in a different fashion than we expected.

The general impression is clearly that *en route*, even within an experimental setting where the users were encouraged to provide feedback to the system, they are concerned with getting to where they are going. This consumes all of their intention and attention to the extent that even frequent reminders about the possibility of rating route segments are ignored or forgotten.

Our attempt at creating a common rating practice with a priori definitions of good, uncomfortable and inaccessible route segments failed. Rather, the users demonstrate that the only consistent type of segment rating produced is that of an inaccessible location. Further, the rating of a segment as inaccessible can be regarded as a *by-product* of the users' selfish goal of solving the navigational task at hand.

In contrast to other community projects concerning user contributed content such as Wikipedia and OpenStreetMap, our users do not constitute a community of practice. Where the builders of content in Wikipedia actively contribute information to reach the common goal of creating the worlds best encyclopedia [7], our users create a common good by being *consumers* of route suggestions, not caring about a community goal of better routes for all.

Two contrasting implications for design stand out as interesting in the future exploration of the OurWay concept: We can hide the route segment ratings entirely from the client application, camouflaging the feedback mechanism as a "request alternative route" feature. This would further distance the user from the workings of the system, and effectively remove the possibility of establishing a community of OurWay users. The alternative is to deploy the system in an *existing* community of practice, where the participants already share a common goal of mapping accessibility, e.g. in public areas and buildings.

6.2 Efficiency

Next, we wanted to know how much it takes from the participants to put the system into a state where it reliably produces good routes:

> How does the number of segment ratings, and the efforts involved in producing them, affect route quality?

It is clear from looking at the ratings provide by users (see Figure 4) that our first two users did the hard work; 22 of the 27 ratings of inaccessible route segments were made by these two users. As our observations reveal, they do not think highly of the route planner the nth time they're directed towards a flight of stairs. In fact, repeated encounters with obstacles of the same type is upsetting even the most enthusiastic of our users. Despite the hard work and number of absolute obstacles met, both participants still report positively about the system's ability to find good routes. One possible explanation is that of our nine users, these are the only two that really experience the system's capability of adapting to user feedback.

Users need to concentrate in order to complete new and demanding tasks [26]. Our users were mostly proficient computer users, many even familiar with advanced mobile phones or PDAs, but they had not previously been confined to a wheelchair. Completing the task of getting around in the wheelchair would thus demand their full concentration, and relegate the handling of the device into the background. We might speculate that proficient wheelchair users would be more accustomed to handling the obstacles thus presented, and that more attention then naturally would be devoted to working with the OurWay application.

On the other hand, to handle the extra load of rating segments, on top of the already challenging task of negotiating physical obstacles in a wheelchair is pointing towards the need for a non-obtrusive and simple interface design. By removing the selection of rating values from the user interface, and replacing it with a button for requesting an alternative route, the perceived effort of contributing to the system would likely diminish.

6.3 User satisfaction

The final research question regarded user satisfaction:

> Do the users perceive the service as useful, and how do they rate route quality?

Satisfaction from the perspective of the tools perceived usefulness should intuitively increase as the number of non-negotiable obstacles decrease. Our debriefing sessions suggests that this happens as expected. However, as our two first users clearly express, they would not have put up with the efforts needed to use the system if this was not in an experiment setting.

This is an indication that to get the number of segment ratings "off the ground", some sort of campaign could be useful, which is perhaps easier to do in an existing community of practice already engaged in accessibility assessment. How well the user knows the building also affects the usefulness ratings: people less familiar with (parts of) the building find the service more valuable.

Interestingly, even when our participants rate the usefulness lower, they tend to think about it as a useful service for others. This contrasts with the way users actually provide feedback to the system, as we have documented. During the navigational tasks, the users are mostly focused on solving the task at hand, with little or no regard for other users and their possible benefit from accumulated segment ratings. However, when retrospectively thinking about the system's usefulness, they soon take the view of the system as a collaborative tool, useful for "the invisible others".

7. CONCLUSION AND FUTURE WORK

We have presented the OurWay prototype, and report from an experiment where we assess the feasibility of the concept in an indoor environment. We are left with two main findings.

First, our previous indications that the concept of collaborative rating of route segments is valid are confirmed through this study by users in a close to real life setting. However, the users established a different segment rating practice than what we suggested to them. In particular, we find that the only predictable kind of rating is given where the user *must* provide feedback in order to get an alternative route from the system. This happens when a non-negotiable obstacle is met.

Second, we have learned that the system works even though the segment ratings are a *by-product* of using the navigational tool. The participants were not engaged in the idea of building a fantastic route planner. They only rarely considered the benefits for other users when they provided feedback on the route segments. They were selfishly driven by the aim of finding the trail through the building there and then. This is indeed an interesting finding, since it shows how the collaborative route system seems to work in a sustainable fashion without a prior community of practice in place. The experimental context, where we encouraged the users to provide feedback to the system, and our failed attempt at building a temporary community of practice strengthens the significance of this finding.

In summary, the concept of users rating accessibility of route segments in navigation systems seems feasible. This insight inspires us to explore the OurWay idea further, and to try it out in a larger setting with more users and diversified user groups. Allowing the technology to be used in a proper real life setting over a longer period of time would give us the opportunity to study usage patterns more closely, and to further investigate the findings we present in this paper.

Since the route segment ratings are mainly collected as a by-product of usage, it would be interesting to consider methods for automatically detecting that a user is diverting from the planned route. Given that there are a number of potential reasons why someone would stop following the suggested route, this is not a trivial problem to solve.

Through our observations, we learned that in most cases, the participants would not be aware of the other previous or future users of the system. Raising the awareness of other users and addressing issues related to trust and reputation in this context is another area suitable for future research.

Measuring the effects of feedback granularity with respect to route quality and willingness to provide segment ratings seems to follow naturally from this study. A formalized software simulation based on our understanding of how and why users rate route segments can be a viable tool for this purpose.

Acknowledgments

We would like to thank Håkon A. Holmstedt for his contributions in the initial stages of the OurWay project, Øyvind Håkestad for contributions to the current prototype and tools, and to the nine users who volunteered to take part in the experiment.

The project is partly funded by the Norwegian Research Council and the Norwegian Mapping Authorities.

8. REFERENCES

[1] Y. Akasaka and T. Onisawa. Pedestrian Navigation System Reflecting Users' Subjectivity and Taste. *Proc. of Int'l Conf. on Control, Automation and Systems*, pages 995–1000, 2003.

[2] Y. Akasaka and T. Onisawa. Individualized Pedestrian Navigation Using Fuzzy Measures and Integrals. *IEEE International Conference on Systems, Man and Cybernetics*, 2:1461–1466, 2005.

[3] W. Balke, W. Kießling, and C. Unbehend. Personalized Services for Mobile Route Planning: A Demonstration. *Proceedings of the 19th International Conference on Data Engineering (ICDE 2003)*, pages 771 – 773, 2003.

[4] L. Beale, K. Field, D. Briggs, P. Picton, and H. Matthews. Mapping for Wheelchair Users: Route Navigation in Urban Spaces. *The Cartographic journal*, 43(1):66–81, 2006.

[5] A. K. Beeharee and A. Steed. A Natural Wayfinding - Exploiting Photos in Pedestrian Navigation Systems. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 81–88, New York, NY, USA, 2006. ACM Press.

[6] N. J. Bidwell, C. Lueg, and J. Axup. The Territory is the Map: Designing Navigational Aids. In *CHINZ '05: Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction*, pages 91–100. ACM Press, 2005.

[7] S. L. Bryant, A. Forte, and A. Bruckman. Becoming wikipedian: transformation of participation in a collaborative online encyclopedia. In *GROUP '05: Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*, pages 1–10, New York, NY, USA, 2005. ACM Press.

[8] K. Cheverst and B. Schmidt-Belz. 4th international workshop on: "HCI in mobile guides". In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, pages 359–360, New York, NY, USA, 2005. ACM.

[9] M. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

[10] K. Z. Haigh, J. R. Shewchuk, and M. M. Veloso. Exploiting Domain Geometry in Analogical Route

Planning. *Journal of Experimental and Theoretical Artificial Intelligence*, 9:509–541, 1997.

[11] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating Collaborative Filtering Recommender Systems. *ACM Trans. Inf. Syst.*, 22(1):5–53, 2004.

[12] H. Holone, G. Misund, and H. Holmstedt. Users Are Doing It For Themselves: Pedestrian Navigation With User Generated Content. In *Proceedings of The First International Conference on Next Generation Mobile Applications, Services and Technologies*, pages 91–99. IEEE, 2007.

[13] C. Huijnen. Mobile tourism and mobile government - an inventory of European projects. Technical report, European Centre for Digital Communication (EC/DC), April 2006.

[14] ISO. 9241: Ergonomic requirements for office work with visual display terminals. *International Organization for Standardization*, 1997.

[15] D. Karimanzira, P. Otto, and J. Wernstedt. Application of machine learning methods to route planning and navigation for disabled people. In *MIC'06: Proceedings of the 25th IASTED international conference on Modeling, indentification, and control*, pages 366–371, Anaheim, CA, USA, January 2006. ACTA Press.

[16] M. Kawabata, R. Nishide, M. Ueda, and S. Ueshima. Graph-based Approach to Context-adaptable PNS and its Application Scenarios. In *Proceedings of the 21st International Conference on Data Engineering Workshops (ICDEW '05)*, Los Alamitos, CA, USA, 2005. IEEE Computer Society.

[17] KDDI. KDDI to launch EZ Navi Walk, a full-scale navigation service for pedestrians. [Online]. Available: `http://www.kddi.com/english/corporate/news_release/2003/1006/index.html` [Accessed February, 2008].

[18] J. Kjeldskov and C. Graham. A Review of Mobile HCI Research Methods. *Human-Computer Interaction With Mobile Devices and Services: 5th International Symposium, Mobile HCI 2003, Udine, Italy, September 8-11, 2003: Proceedings*, 2003.

[19] J. Kjeldskov, M. Skov, B. Als, and R. Høegh. Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. *Proceedings of Mobile HCI 2004*, 2004.

[20] E. L.-C. Law, E. T. Hvannberg, and G. Cockton. The Future of Usability Evaluation: Increasing Impact on Value. In *Maturing Usability*, pages 344–378. Springer, 2008.

[21] L. McGinty and B. Smyth. Shared Experiences in Personalized Route Planning. In S. M. Haller and G. Simmons, editors, *FLAIRS Conference*, pages 111–115. AAAI Press, 2002.

[22] N. McNamara and J. Kirakowski. Defining usability: quality of use or quality of experience? *Professional Communication Conference, 2005. IPCC 2005. Proceedings. International*, pages 200–204, 10-13 July 2005.

[23] R. Priedhorsky, B. Jordan, and L. Terveen. How a personalized geowiki can help bicyclists share information more effectively. In *WikiSym '07: Proceedings of the 2007 international symposium on Wikis*, pages 93–98, New York, NY, USA, 2007. ACM.

[24] S. Rogers and P. Langley. Personalized Driving Route Recommendations. In *Proceedings of the AAAI Workshop on Recommender Systems*, January 1998.

[25] N. Ronald, L. Sterling, and M. Kirley. A conceptual framework for specifying and developing pedestrian models. In *Proceedings of the 16th Biennial Conference of The Modelling and Simulation Society of Australia and New Zealand (MODSIM 2005)*, 2005.

[26] B. Shneiderman and B. B. Bederson. Maintaining concentration to achieve task completion. In *DUX '05: Proceedings of the 2005 conference on Designing for User eXperience*, page 9, New York, NY, USA, 2005. AIGA: American Institute of Graphic Arts.

[27] A. Sobek and H. Miller. U-Access: A web-based system for routing pedestrians of differing abilities. *Journal of Geographical Systems*, 8(3):269–287, 2006.

[28] Telecom Tribune. NTT DoCoMo To Launch Navigation Service for Pedestrians Using Enhanced GPS Technology. *Telecom Tribune*, 14(10):2 – 2, January 2000.

[29] M. Wuersch and D. Caduff. Refined Route Instructions Using Topological Stages of Closeness. In *Proceedings of Web and Wireless Geographical Information Systems, 5th International Workshop*, volume 3833 of *Lecture Notes in Computer Science*, Lausanne, Switzerland, December 2005. Springer.