Enhancing In-Car Navigation Systems with Personal Experience

Benjamin B. Bederson
Human-Computer Interaction Lab
Computer Science Department
Institute for Advanced Computer Studies
3171 A.V. Williams Building
University of Maryland
College Park, MD 20742
+1 301 405 2764
bederson@cs.umd.edu

Aaron Clamage
Human-Computer Interaction Lab
Computer Science Department
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742

Catherine Plaisant
Human-Computer Interaction Lab
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742
+1 301 405 2768
plaisant@cs.umd.edu

Number of words in abstract, text and references: 6774
Number of figures: 6
Total effective # words = 6,820 + (6 * 250) = 8,274

ABSTRACT

Computers are extremely powerful for data processing, but less adept at handling problems that involve subjective reasoning. People, on the other hand, are very good at these kinds of problems. We present a framework for adding subjective human experience to in-car navigation systems. People often rely on their own experience when planning trips, choosing the route that seemed fastest in the past, the one that was the prettiest, or the one recommended by a friend. This led us to develop a set of methods to help people record their personal driving history, add textual annotations describing subjective experiences, and share their data with friends and family, or even the broader community. Users can then learn from their own data, or harness the multiplicity of individual experiences to enjoy new routes. This approach can be used in conjunction with traditional in-car navigation systems.

INTRODUCTION

Designers have just begun to realize the potential of using human power to accomplish tasks that computers, automated sensors and signal analysis techniques cannot do in a timely and accurate fashion. For example, researchers at Carnegie Mellon have created online games to attract users to annotate images or add facts to a knowledge-base [10], [11]. Google Co-op invites people to manually annotate web pages to improve Google's search capability. And search engines work so well in part because they analyze the result of human activity such as creating structured documents and linking among them.

Yet, most current navigation solutions rely entirely on objective, automatically collected data, ignoring human subjectivity. Personal experience plays a huge role in route choice. Traffic reports and websites do not always provide relevant, timely and accurate information and drivers often rely on their own knowledge – or impressions - when planning trips to the airport, train station, friends, family, etc. Most use a trial and error approach [6]. People have different criteria for the "best" route: fastest, most scenic, lowest fuel consumption, or even perceived safety - so combining objective and subjective measures may enhance the route selection process.

Our approach differs from other route planning systems by starting with a decentralized data collection process (Figure 1). By recording their own driving history, people can learn about the routes they actually use. By adding simple textual annotations to those routes, they will later benefit from data that cannot be recorded automatically, such as the beauty of the scenery or their feelings of insecurity. Aggregating personal data has the potential to aid in the discovery of patterns and confirm (or not) impressions drivers have about the characteristics of routes. While all users get immediate benefit from collecting and analyzing their personal data, some will reap greater benefits from sharing with friends and family, and then may choose to fully anonymize the aggregated data, share it with larger communities, and access more route data in return.

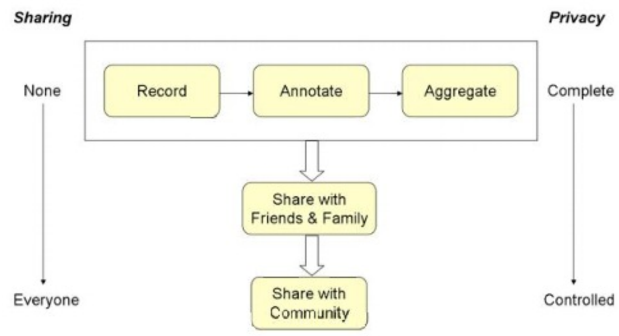


Figure 1: Personal experience aggregation and sharing. Drivers record, annotate and aggregate their personal data, and can choose to share with friends, or a larger community.

In this paper, we present a design framework for collecting, aggregating and sharing personal driving history as a case study of taking advantage of untapped human experience. We begin by discussing a survey we conducted to determine the type of questions people have when selecting routes, and how willing people are to share their personal driving experiences. We then discuss several scenarios using an early prototype to illustrate the recording and viewing of routes and mock-ups to show how annotations and sharing might take place. We also discuss methods of maintaining accuracy and anonymity as well as techniques for aggregating large collections of experiences and presenting that information to users in a meaningful way.

RELATED WORK

Harnessing Human Experience

Existing systems to collect human experiences can be grouped using a simple taxonomy (Figure 2). Some of these systems aggregate data, while others do not. Either way, they motivate people to participate in four basic ways. Sometimes, there is no extra effort involved. Many e-commerce sites, for example, track when users purchase items to make recommendations to others. People may be motivated to spend some effort because it is fun (e.g. they play a game). Or, they may get some direct personal benefit from participating. Finally, people are often motivated if they feel they can help make the world a better place. Our work will focus on the last two kinds of motivation – personal benefit and altruism – engaging people because they want to be involved.

Figure 2: Taxonomy of existing systems, using 4 levels of motivation to participate and 2 levels of aggregation.

Some of the best known experience collection systems are websites like www.amazon.com, www.netflix.com, and www.ebay.com, which use collaborative filtering. Amazon lets people write reviews and rate items from one to five stars. The ratings are aggregated while the reviews are simply listed individually. Readers can also rate the reviewers. People can then sort the individual reviews by those that are found to be the "most useful." Other services like Netflix recognize that many users favor restricting access to their comments and recommendations to a smaller circle of friends. Systems like eBay collect reviews in order to control malicious use. A seller's average rating may influence whether or not he gets more business in the future. People are willing to invest some effort to help the greater good, and because they benefit personally when other people give positive feedback about them.

Many systems appeal to our sense of altruism. Google Co-op invites users to manually annotate web pages to improve Google's search capability. Large fact-gathering communities, like Wikipedia and many web forums, explicitly collect intellectual contributions from users. These efforts assume that a large number of people will volunteer to contribute if they feel they can help make the world a better place.

Other systems motivate people to spend effort simply because it is fun to do so. Researchers at Carnegie Mellon have built games that can significantly improve computerized tools by enticing users to contribute knowledge. For example, to improve image search, Von Ahn and Dabbish introduced the ESP game [10] to collect and evaluate metadata about images. Building on their success, von Ahn *et al*. later developed Peekaboom [11], a game to determine the locations of objects in images.

Social networking websites such as www.myspace.com, www.linkedin.com, and www.facebook.com have built networks of millions of people which include personal experiences, but do not analyze the data or aggregate it in meaningful ways. Systems like Google's page-rank, which estimates importance of web pages based on others linking to them, utilize efforts previously spent. Other sites, like Yahoo Answers and Amazon's Mechanical Turk, connect individual question asker's with answerers, taking advantage of broad participation, but still do not aggregate the data or build structure from content.

Route Planning Domain

For many years, researchers have developed systems that observe individual drivers to estimate road speeds. Cameras and inductive loop detectors embedded in the pavement routinely produce information about the speed of traffic, but the data is sparse and requires significant infrastructure. In the late 1990s, the Universities of Maryland and Virginia and several cellular companies pioneered efforts to use location data from cell phones to estimate travel time [9]. Since then, countless studies have examined different ways to do this, using statistical techniques, artificial intelligence, and pattern matching [2]. Fawcett and Robinson estimate road speeds for small time slices to predict the optimal route at a given time [3]. Harrington and Cahill tag time records with weather and road condition information to predict travel times by context-matching [5]. These efforts, however, tend to focus on improving traffic management and reducing traffic jams. They do not help users understand their personal experiences, aggregate those experiences or add other useful information.

Some researchers have looked more closely at personal route planning. Konishi *et al*. [6] described a scenario for selecting the best route by aggregating driving history and providing summary information, but there was no notion of annotation or subjective route attributes, and no interface was built. Letchner *et al*. developed the "TRIP" system [7] to incorporate time-variant traffic data and user preferences into route planning. They collect GPS traces from a large number of drivers to estimate traffic speeds for small time slices, taking user preferences into account by suggesting routes similar to a driver's previous trips. However, they do not let users explicitly add information to the system or discuss how to combine data from multiple drivers.

SURVEY AND USER NEEDS

In order to inform our theories and get a better picture of what a personal experience collection system might look like, we started by conducting a web survey to find out what people think about their current driving activities and needs (www.cs.umd.edu/hcil/routelens/survey). The survey asked questions about their current driving and route finding habits. It focused on repeated routes, the criteria people used for choosing those routes, and the information sources people used for determining which routes to take. It also asked about the relative importance of average route speed vs. consistency. Finally, it asked about participants' willingness to share data about their driving and route selection behaviors and ended with some basic demographic questions.

Terminology

We define some basic domain terms for use in the survey. We use the word *route* to mean a particular path or set of roads from one location to another. Each time a user drives along a route, we call that a *trip*. The term *route collection* is used to designate all the different ways to get from one location to another. Finally, *locations* are the named end points of a route.

Respondents

We solicited participation for the survey by sending an email to our colleagues and asking them to forward that email to anyone they knew. In all, there were 292 respondents, covering various fields, including computer science, IT, software design, and health care. Roughly 67% were female, ranging in age from 18 to 75, with 53% between the ages of 46 and 60, and 92% from the United States. Other regions represented included Asia, Australia, Europe, the Middle-East and Canada.

While the participants of this survey represent a reasonable diversity of age and profession, it should be noted that this was a self-selected sample, it is relatively small, and is not representative (statistically speaking) of all drivers in the United States or elsewhere. But, it is taken from a group of technology savvy users that are more likely to use this sort of approach. So, while we believe this data is broad enough to be informative, it is not definitive. But since our goal is to show only that at least some non-trivial set of drivers would be interested in subjective data for route finding, we feel that this survey meets those needs. It is not our goal to show that this approach, for example, would support a viable business today – only that it is interesting and worth continued and deeper understanding.

Domain Analysis

Many respondents admitted they did not know the best route for their common trips, with nearly 80% agreeing that a system to track personal driving history would be useful. One participant even commented “the number of reasons this would be useful to so many people are too numerous to count here.” Nearly 60% of respondents said they argue at least once in a while about which route is best. About 80% could think of at least three trips with alternate routes that they drive regularly, with 33% listing more than five. Finally, most respondents said they usually choose routes by considering alternatives and selecting the one that seems best, apparently by using their own judgment to make a decision based on all the available information. As we suspected, it seems that many people think personal routing is an important problem and they often use their own experience to make better route choices.

Questions and Tasks

Choosing the right problem domain is not enough to build a successful personal experience aggregation system. To motivate people to participate, the tasks supported by the system have to be personally beneficial to them. We had users rate how useful our initial tasks were, as well as suggest their own tasks.

Our final task list can be classified into several groups – personal, collaborative, and other. Personal tasks are those that use an individual’s personal driving history alone. Collaborative tasks are those that involve information from other users. Finally, other tasks involve some external information. The following are examples of personal tasks:

- Find best route for given trip and departure time
- Find best departure time for a given trip
- Find routes to places I have been, but do not remember exactly where they were
- View summary information for my routes (e.g. average speed, number of stops, total time)
- Find a route based on my personal driving style
- Add my own information to routes (e.g. road conditions, scenery, safety, points of interest)
- Find a route based on my annotations (e.g. minimize highways, maximize safety)
- Find points of interest near a route
- Find route near points of interest
- View map showing only roads I take frequently

Of course, many of these tasks could also be performed with data from family and friends, or other community users. However, some tasks are more suited for collaboration than others. The following are examples of collaborative tasks:

- Get better statistics by sharing routes with others
- View others' routes to find a good alternative
- View others' annotations to their routes
- Determine the validity of others’ annotations
- Find routes to places I have not been before
- Find routes similar to a particular route I like
- Find people with similar routes for carpooling
- Compare differences between drivers
- Generate detailed directions to share with others
- Provide interactive communication between drivers

Finally some tasks are neither personal nor collaborative. For example, twelve respondents independently commented that they would want to view real-time traffic information to choose alternate routes based on exceptional circumstances, like accidents or construction work. While, this information might come from annotations, it would more likely be drawn from external resources maintained by transportation authorities, but customized to their personal routes.

Route Criteria

When collecting personal experiences, it is also important to distinguish between data that can be tracked automatically and data that has to be manually added to the system by users. The most common task above was to find the best route based on some user-specified criteria. Thus, it is important to know what kind of data people care about when choosing a route. We asked respondents to rank the importance of several route criteria and suggest their own. Results show that while people care a lot about the measurable attributes of a route, they also care about many more subjective attributes. Examples include:

- Distance
- Travel time
- Time stopped
- Number of stops
- Driving speed

Other attributes deemed important cannot be easily recorded automatically, and need to be added as annotations. Examples of these kinds of attributes include the following:

- Scenery
- Safety
- Fuel consumption
- Road conditions (e.g. potholes)
- Nearby points of interest (e.g. cheap gas station)
- Types of roads (e.g. back-roads, busy highways)

Some of these attributes may become automatable, e.g. fuel consumption could be calculated if the route tracking system were integrated with a car's on-board computer. However, there will always be subjective attributes that cannot be measured - like beauty or feelings of safety - which our survey indicates are very important.

Sharing

Finally, we wanted to know if people were willing to share their personal data with others. This is crucial if we want to collect and aggregate individual driving experiences. Surprisingly, most respondents (~75%) said they would be willing to share some or even all of their data. Privacy is of course an issue, with one respondent stating "I would not want criminals to have access to when I would not be home."

Most people wanted the ability to choose who to share with and whether or not to remain anonymous. About 35% said they would share with a group they could choose, and nearly half said they would share with anyone if they could hide their identity.

Respondents also wanted to control which data they shared with others. Most preferred to choose a subset of their favorite routes, annotations, and statistics, with one participant saying "I would share any information that I did not classify as personal." Seventeen respondents independently commented that they would want to clip the end points of their routes, so others could not see where they came from, or where they were going exactly. Many people were hesitant to tell everyone their tricks. One person said "I would not share too many of my secret methods … If everyone knew the exit chute, then it would not be quick." Another noted "Sharing my secrets for alternate routes may adversely affect that route for me in the future – Selfish probably, but a survival technique in a heavily congested area." People, however, wanted to share these tricks with their friends and family.

Finally comments suggested that people were aware of the public benefit of sharing elements of personal experiences (e.g. road condition annotations or wait times at intersections), and had some altruistic motives to share their knowledge of the routes (e.g. with local history information).

In summary our survey indicates that people are willing to share if they can reap some personal benefit, help their friends and family, or even help the greater good, but they need control over what to share, who to share with and whether or not to remain anonymous.

A DRIVER CENTRIC APPROACH

Following the personal experience aggregation process outlined earlier, and informed by our survey, we designed a set of methods and interaction scenarios to motivate individuals to collect their personal driving experiences and share those experiences with others. We built an early interface prototype, called RouteLens, which we use to illustrate how drivers might record, aggregate and learn from their personal data. For annotation and sharing, mock-ups are presented.

Decentralized Data Collection

Current efforts to collect traffic information are typically initiated by transportation authorities, using sensors embedded in the road, cameras or cell phone information, relying on a great deal of supporting infrastructure. Soon some cars will also automatically collect and report on road conditions (e.g. speed, rain and ice etc.) In contrast, our approach utilizes personal information collected by the drivers themselves on personal devices, maintaining complete privacy control and requiring no infrastructure, aside from GPS services. Each time an individual drives along a route, her personal mobile device - if purposefully turned on – will track the location and compute distance, time, and speed of the trip (Figure 3). Note that the actual tracking of GPS to create valid road-based routes is a well-studied problem, but is beyond the scope of this paper and will not be discussed here.

Figure 3: Mobile interface. Simple buttons start and stop recording, labels display position, and map shows route (updating route view only partially implemented)

Personal Data Analysis

View routes

Once users have recorded some data, they can view all their destinations and routes in a desktop visualization (Figure 4). Our experience suggests that even viewing the shape of well known routes can be useful and surprising.

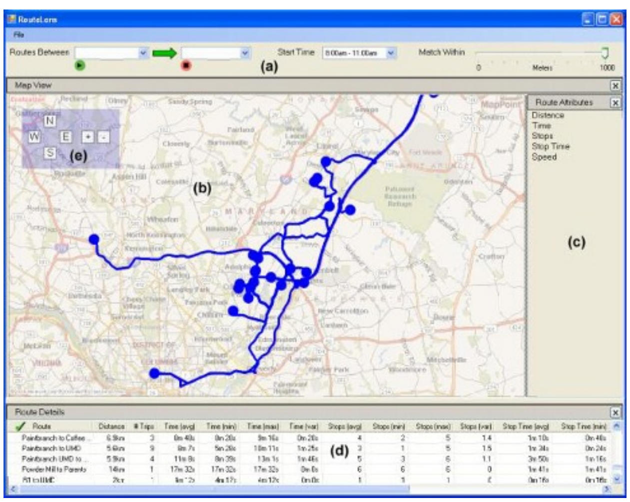


Figure 4: Desktop interface consists of components to (a) select routes, (b) browse data, (c) compare routes, (d) view detailed statistics, and (e) zoom and pan.

Since the route information is personal, users get other immediate benefits, even before aggregation or adding a single annotation. Consider this scenario:

John is going to a doctor's office, where he has not been for months. He knows the general area of that office, but does not remember the exact location. He asks to see where he has driven in the area and immediately recognizes the doctor's office location.

Uploading Multiple Trips

In order to aggregate statistics properly, the system needs to determine if each trip recorded by the mobile device was driven along an existing route, or a new route altogether. This could be done completely automatically with pattern matching techniques, but the result is not accurate. Instead, the system makes a suggestion for each trip, and lets users accept or correct those suggestions. When uploading trips, the interface suggests an existing route if there is a close match. The new trip appears dotted orange, and the suggested route blue, making visual matching easier. Users can hit a key to add the trip to the suggested route, select a variant route, or create a new route altogether. Practically, large numbers of trips can be entered in a matter of seconds once common locations and routes have been uploaded. If a "secret" trip or location had been recorded by mistake, it can be erased. Our impression is that most people do not seem to have locations to hide, and the mobile device can simply be left off. Stops along routes (e.g. for gas) could be removed automatically or the whole trip can be ignored.

Learning from aggregation

Once users have collected enough trips, they can review summary information to compare differences in distance, travel time, number of stops, total stopped time and speed. Locations are selected and when an attribute is chosen, the map highlights the best route (Figure 5). Users can also make more subtle comparisons using variance:

Mark frequently drives to the coffee shop and wants to reduce his time stuck at traffic lights. He has used three alternative routes and selects "Stop Time" to compare them. He sees that route the route that takes 495 to the coffee shop has the shortest average stop time, but the most variance. On the other hand, the route that takes Paintbranch Road has only a few more seconds of stopped time on average, and less variance.

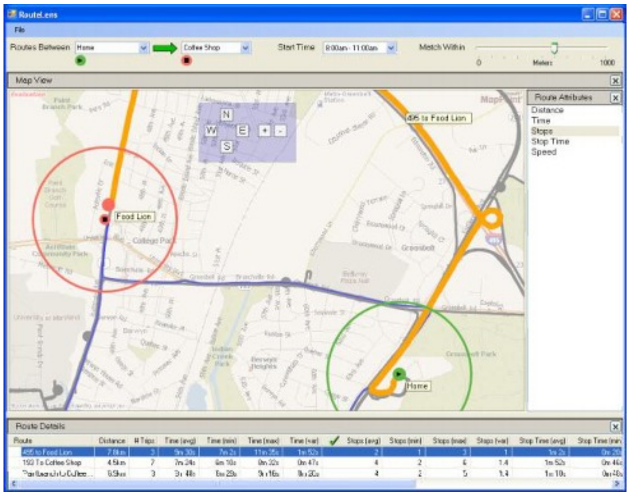


Figure 5: Attribute queries. Users selected "number of stops" as the attribute to compare routes between 2 locations. The best route is shown with a thick orange line.

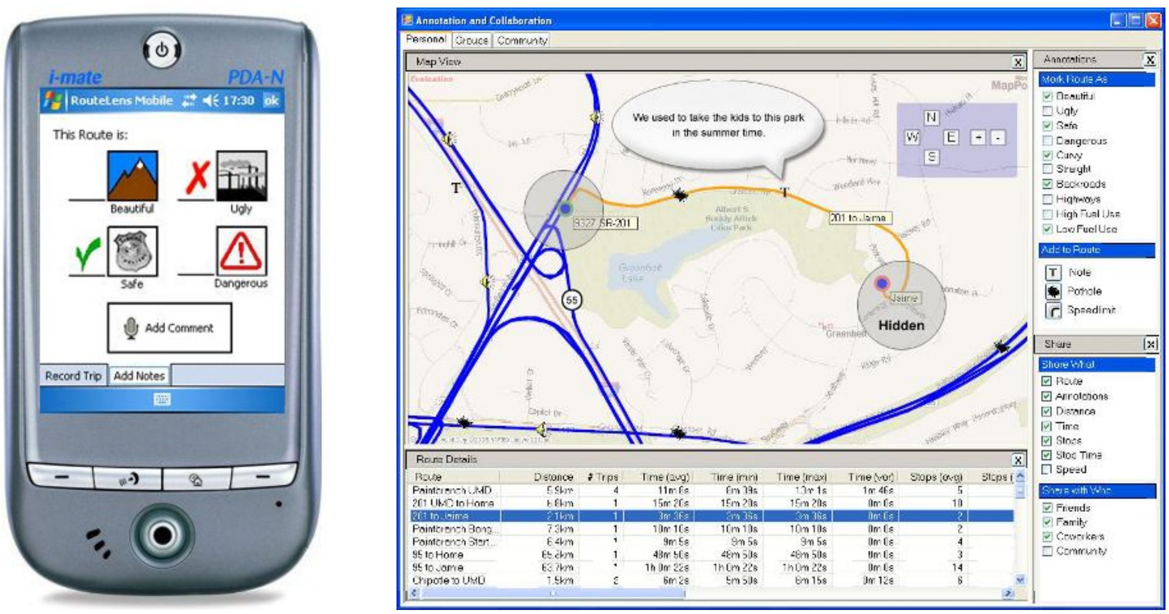
Other Information

One of the popular tasks from our survey was to view real-time accident information. While personal data (e.g. road conditions, location of snow drifts, down trees) might help answer this question. The desktop interface adds real-time traffic incidents from public sources that are located close to their personal routes. Potentially statistics about accidents could also be added to the route information.

However, the display of real-time accident information as well as any other kind of data display that would be useful in the car for use while driving brings up the crucial issue of safety. As with GPS data interpretation, we recognize that the safety issues are very important – but they have been studied extensively elsewhere, and are beyond the scope of this paper since the kinds of solutions offered for other in-car systems will apply to the subjective kind of data suggested in this paper as well.

Annotation

While annotations are not required for the system to be useful, our survey suggest that adding annotations will increase benefits to users. First, users want to compare routes based on more subjective criteria, like the beauty of the scenery, the ease of driving, or the safety of the neighborhood.



(a) (b)

Figure 6: Adding annotations. In (a), drivers – or better, passengers – can add structured notes and voice comments on the mobile device while en route. In (b), users add structured notes and free text after traveling, edit annotations, and share their data.

By adding these details to the data themselves, drivers will benefit from having the information they actually care about:

Wendy needs to drive to the train station and knows several routes. Viewing her annotations, she sees one route has "cheap gas" and a "fast decent Mexican restaurant." Figuring she can stop for a quick bite and save some money, Allison chooses that route.

People might also record annotations to help them remember important information they want to share with friends or family. Consider the situation of explaining directions to a friend:

Jim is driving to the Greenbelt Community Center, where he will meet his friends for an art exhibit next week. While approaching a tricky turn he knows well, he records a note: "Curve right onto the exit ramp for the *parkway, but as the ramp straightens avoid exiting by merging one lane to the left."*

There are several other reasons to record annotations as well. People may want to share their knowledge of local sites: "That green house is the oldest one in College Park;" record personal memories with family: "This is the hill where we took the kids sledding each year," or share important notes with the community: "I wouldn't drive here at night."

Classification

The annotations we describe here can be classified along several dimensions, including *type*, *level*, and *localization*. We identify two basic types of annotation: *structured* and *unstructured*. For some route attributes, like safety, scenery, or road type, it makes sense to provide a structured way for users to quickly rate or categorize their routes. While these kinds of annotations are easy to aggregate, they are limiting. Users who enjoy annotating might add less structured notes, involving voice, pictures, sound or free text:

After passing an area with dozens of trees down, Jerry presses the "Add Comment" button and says "Whoa, check the damage of the July 4th tornado." It will be replayed to his friends and family ahead of the location when they drive in the area (and choose to hear annotations.)

We further identify two levels of annotation: *en route* (Figure 6a) and *after traveling* (Figure 6b). The most natural time to add notes is often while en route. Like any device widely used in the car (phones, PDA's or direction finders) interaction while driving should be limited, done by a passenger or when stopped. While at home, users might add or edit annotations, e.g. adjust locations of comments or add more detailed information to give directions to others.

Finally, annotations may have different degrees of localization, pertaining to an entire route: *"Always lots of trucks";* a segment of a route: "*Last mile is unpaved*" or a specific point along a route: "go *35 - Speed trap* ".

Presentation

Just as there are multiple ways for users to record annotations, there are multiple ways to view and interact with recorded data. People might view their annotations before driving, with their home or office computer, or on the road, with their mobile device. The presentation will need to be different in each case and will depend on how localized the data is. Users might listen to general route annotations at home, but want to be notified while driving if they are approaching a tricky turn or a pothole. The mobile interface will have to be very restricted, due to safety concerns and limited screen space. While driving, users should only view data and listen to or add voice annotations, but when the car is stopped, or if a passenger is present, people can perform more complex tasks.

Sharing

From our survey, we know that people are willing to share, but privacy is a key issue. Our solution is to make collaboration a multi-leveled, opt-in task, letting people choose who to share with, what to share, and whether or not to remain anonymous. The system can guide them through a natural progression, from individual to community use. Drivers will benefit from recording their personal information alone, but the more they are willing to share, the greater the benefit they will receive. They may start by interacting with just friends and family, and eventually discover they can learn more by collaborating with the rest of the community.

Who to Share with

We suspect that some people are more likely to ask their friends and family for advice than consult public data repositories when planning unusual trips, because of trust and shared knowledge of driving preferences. Recent work related to social network analysis shows that including information about the trustability of data sources can increase not only the perceived, but the actual quality of the information for a particular individual [4]. An example usage is:

Mary needs to drive to Annapolis for a concert after work. She could MapQuest it, but instead she trusts that her coworker Anne, who commutes from Annapolis every day, will know the best route. Anne sends her recommended route and share her historical data so Mary can decides when to leave.

A driver might also want to leave messages only to his spouse and kids, like "*Honey, monster pothole in the intersection*"; or ask the annotations from neighbor who has unique knowledge of landmarks:

Jane is driving with out-of-town family but can never remember which buildings are significant. She has turned-on the commentaries of her friend Isabelle, an architect-historian. As they pass by an old house, everyone hears: "Montpelier Mansion, a good example of Georgian architecture that has hosted many famous people, from George Washington to Abigail Adams."

Sharing route data with friends and family alone is useful and give good control over privacy, drivers will also benefit from data aggregated from large numbers of users.

Sara is meeting a friend for dinner at a restaurant she has never been to. She searches the community data, telling the system to "minimize highways," since she prefers back-roads, and to "maximize safety" to avoid driving through any dangerous neighborhoods. Finally, she sorts the results by "number of ratings" to choose the route with the highest confidence level.

When more users share annotations with the community, public benefit increases. Transportation agencies can use aggregated reports of road conditions to spot problems. During an emergency, thousands of users could decide to share their reports of the conditions at their locations (e.g. down trees, power outages, fires), therefore dramatically increasing the information available to first responders.

What to Share

Drivers may not want to share all of their routes, statistics and annotations. If someone has secret tricks like where to change lanes, he will only share that information with friends. People usually do not want to give away personal information or reveal their identity (even though there are many exceptions). It is important to provide mechanisms to allow users to limit what they share, or to share data in an anonymized way:

John wants to share his favorite routes and recorded statistics with his small town community, but does not want people to see how fast he drives or where he goes exactly. He chooses to share everything but removes travel speeds and clips all end points of the route to hide them.

Anonymity

Drivers may also want to hide their identity. They should be allowed to choose whether to remain completely anonymous, show some personal information, or reveal who they are. Again, the more information they share, the greater the benefit they will receive. For example, revealing personal information may help users identify people they have something in common with:

Bob is very anxious about privacy. He shares nothing at first… Later on he decides to share aggregated data about route segments (i.e. between 2 intersections), never even revealing entire routes, to access similar data from others.

Harry spills details of his life on his blog and loves to share all his notes on the routes he uses. He now wants to find someone to carpool with. He enters start and end locations and preferred departure time, to search for people with a similar commute. After sorting the results by users with similar driving style and schedule to his, a message is sent to the few people who chose to identify themselves to their neighbors.

These two scenarios bring up one challenge with this approach which is the statistical validity of the aggregated data. Given that individual users are not likely to generate much subjective data, it is by its nature, going to be idiosyncratic. Surprisingly, this does not invalidate the utility of the data. If individuals are sharing with other known individuals, then the personal and idiosyncratic nature of the data is precisely what is going to make the data trustworthy. It is only when the data contributors are not as well known by the users that statistical validity will be more valuable. However, as there are more contributors so the data becomes less personal, the data becomes more statistically aggregatable. So, these two potentially conflicting factors (trustworthiness of individuals vs. reliability of data) are likely work hand-in-hand with different use cases (i.e., personal recommendations vs. collective knowledge).

Accuracy

If participants are self-selected, and the data self-contributed, there may be issues of selection and content bias, which could be unintentional or malicious. We propose several methods to mitigate these problems:

- Minimize impact by having many participants
- Detect and discard outliers
- Community ratings of users and content
- Automatic detection of personal and group consistency

People's perceptions of subjective attributes vary greatly. What one person considers frightening, another might find exciting. Therefore, collaborative filtering can provide a way to focus on data from people with similar sensibilities.

Presentation

Users should be able to search any subset of their personal and shared data to find a good route. As shown in Figure 6b, they can view routes, statistics, and annotations from other groups they share with or from the broader community. For the aggregated community data, the interface would show an overview first, and let users filter down by entering criteria they care about.

Aggregation

Aggregating different kinds of data from potentially thousands of users is particularly challenging. We suggest a range of approaches.

For the route data and recorded statistics, the process should be fairly straight-forward. Once many users share their routes, there will be a lot of overlap. One user may drive along a route from his house to the coffee shop that takes 495, while his neighbor may have a route to the Chinese Restaurant, next to the coffee shop, which also takes 495. There is probably no need to show both of these routes in the community data. We could use pattern matching techniques to find similar routes and then combine the statistics to get more accurate summary information. Once we find aggregated routes, aggregating structured annotations (e.g. safety, beauty) is fairly straightforward.

Unstructured annotations, such as free text, are more difficult to aggregate. Linguistic analysis techniques could be used to group similar notes, by searching for keywords in context. For example, Subramanian *et al.* built the Opinion Analysis System (OASYS), to track world-wide opinions on matters of national security by searching news feeds for positive or negative keywords [1]. Approaches like this, however, can be problematic since there are many ways to phrase a similar statement, such as "This route is really dangerous," and "I drove through several bad neighborhoods." To improve aggregation, human skills could be used to supplement computational methods.

CONCLUSIONS

We have presented a framework for harnessing human experience to help individual drivers make better route choices. The framework was motivated by a survey and developed by implementing a partial prototype and mockups. Our approach differs from other route planning systems by focusing on subjective human experience. We have shown how driver experiences can be aggregated from the ground up, first by helping people record personal history, then by providing them with a mean to share data with family and friends, or even the broader community. We discussed issues of maintaining anonymity and accuracy, the technical challenges of aggregation, and how to present the data to the user.

We conducted a survey that, while limited in generality, collects data from a self-selected sample of individuals that we believe represent those technically savvy users that are most likely to be initially interested in the sort of system we propose. Among the survey participants, we heard that they care a lot about subjective route criteria, while previous work clearly indicates that they can also benefit from automatically recorded information. The survey also indicates that they want to have access to external data, like real-time traffic information. This leads us to believe that the best system would provide a combination, of automatic, annotated, and other supplemental information, when available

Our survey also indicates that many people are willing to share everything with friends and family and some things with the community. We believe this provides strong evidence that experience aggregation systems would be useful. Even if people only share their data with a circle of friends, they will reap significant benefits. If they share anything with the broader community, the overall benefit to everyone will be even greater.

In general the challenge of personal experience aggregation systems seem to be to provide:

1) immediate benefit even without sharing,

2) increased benefit thru annotating and sharing with a selected few users,

3) the potential to serve the greater good with global sharing of anonymous information.

However, it is the nature of this kind of system that there is a "chicken and egg" problem in deploying it in that it doesn't become useful until there are a lot of participants, but people are less likely to participate until it provides value to them. However, this shouldn't scare us off too much as there are plenty of systems whose primary value comes from "network effects" (such as the very popular social network websites) which have overcome this problem. If individually contributed subjective data can provide value when a lot of people participate, then it is worth pursuing how to motivate them to do so.

This paper describes a framework for collecting individual subjective data. But there is a lot more to study to ensure the excellence of such a system. Future work might include a more complete implementation of our route planning system, or exploring other areas application domains. We could also envision developing a general infrastructure to support a set of common tasks. We hope that the design investigation reported in this paper will inspire others to collect, aggregate, and communicate human experiences.

ACKNOWLEDGEMENTS

We wish to thank John SanGiovanni for his input and support, Ben Shneiderman for contributing the original idea of recording and comparing routes, and Makoto Konishi for his original development of that idea. This work was supported in part by Microsoft Research.

REFRENCES

1. Cesarano, C., Dorr., B., Picariello, A., Reforgiato, D., Sagoff, A. and Subrahmanian, V.S. Opinion analysis in Document Databases. In the *Proceedings of the AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs*, AAAI Press, 2006, 21-26.

2. Chung, E. Classification of Traffic Pattern. In the *Proceedings of the 10th ITS World Congress, 2003.*

3. Fawcett, J., and Robinson P., Adaptive Routing for Road Traffic. IEEE Computer Graphics and Applications, 2000. **20**(3): 46-53.

4. Golbeck, J., Parsia, B., and Hendler, J. (2003). "Trust Networks on the Semantic Web" in *Cooperative Information Agents VII*, Springer Berlin / Heidelberg, 238-249.

5. Harrington, A., Cahill, V. Route Profiling: Putting Context to Work. In the *Proceedings of the Symposium on Applied Computing (SAC 2004)*, ACM Press, 2004, 1567-1573.

6. Konishi, M., Plaisant, C., Shneiderman, B. Enabling Commuters to Find the Best Route: An Interface for Analyzing Driving History Logs. In *Proceedings of the IFIP Conference on Human-Computer Interaction (Interact2001)*, IFIP IOS Press, 2001, 799-800.

7. Letchner, J., Krumm, J., Horvitz, E. Trip Router with Individualized Preferences: Incorporating

Personalization into Route Planning. In *Proceedings of the Conference on Innovative Applications of Artificial Intelligence (IAAI-06)*, AAAI Press, 2006.

8. Longman, P. American Gridlock. *US News and World Report*, May 28th, 2001.

9. University of Virginia Center for Transportation Studies, Virginia Transportation Research Council. Probe-based Traffic Monitoring State-of-the-Practice Report. NCHRP Project 70-01, for Private-Sector Provision of Congestion Data, 2005

10. von Ahn, L., and Dabbish, L. Labeling Images with a Computer Game. In *Proceedings of the SIGCHI conference on Human factors in Computing Systems (CHI 2004)*, ACM Press, 2004, 319-326.

11. von Ahn, L., Liu, R., Blum, M. Peekaboom: A Game for Locating Objects in Images. In *Proceedings of the SIGCHI conference on Human factors in Computing Systems (CHI 2006)*, ACM Press, 2006, 55-64.