# **Planning and Prediction: Learning Driver Preferences**

Anind K. Dey (anind@cs.cmu.edu) and Drew Bagnell (bagnell@ri.cmu.edu) 412-287-9620

## **Project summary**

The focus of our proposed project is to support drivers in the Oakland-downtown corridor (and other drivers in Allegheny County). Drivers know multiple routes to get between Oakland and downtown, and use their best judgment to select one to take, based mostly on time of day and day of week. They often do not have access to the information they need to make an informed choice about which route to take. The goal our work is to learn the preferences of drivers, use this information to predict the route and destination of each driver, and then proactively provide relevant traffic, weather, event information that will likely impact the driver's commute.

### Problem definition and desired results

In the past decade, computer systems that interact with people by providing routes, usually for vehicles, have become a crucial part of our lives. Businesses depend on systems like *Microsoft's MapPoint* to plan routes for transportation and delivery, drivers depend on web-based interfaces like *MapQuest* and in-car navigation systems like *Garmin's Nuvi* (Figure 1) to obtain accurate driving directions, and even walking and public transportation routes have come under the purview of such systems. These tools build on intelligent computing algorithms, like A\* search [Hart68], to provide efficient routes for users. While well developed in terms of reliable implementation and availability, such systems suffer important limitations apparent to users.



Figure 1. A typical modern in-car navigation system.

First, the route costs in these systems are handcrafted functions of manually annotated descriptions of the road network. These annotated descriptions are very costly to obtain [CNN06], and often the features of the descriptions and handcrafted functions do not capture the desirability of a particular route in the various dynamic situations that affect drivers' preferences, such as construction, time of day (e.g., rush hours, night), and different weather conditions. The full information on which drivers base their route planning decisions can be extremely difficult to capture or require a handcrafted function to represent. For example, choosing a route after a winter storm may depend on the grades of different road segments and possibly the municipality's salting and plowing policies for those segments as well.

Second, routing systems are nearly entirely ignorant of the unique preferences of a particular user when filling a request for a route. For instance, users more familiar with a particular neighborhood, such as the downtown-Oakland corridor, will welcome more complicated directions while novice travelers will be willing to pay a moderate cost of time for a set of directions that is simple to remember and execute. In interviews with 10 drivers and surveys of 21 others, we found that some users preferentially avoid difficult to negotiate traffic situations, multi-lane highways, toll roads, or roads with a high grade (particularly in winter conditions). There were individual differences among our subjects, meaning that

there is not a single optimal route between a given origin and destination that would satisfy all drivers. Thus, it would be almost impossible to construct simple rules or heuristics by hand for predicting a driver's most probable or preferred path.

Finally, and most importantly, modern navigation systems rely upon an interaction model where a user always requests a route to a destination through some interface. This is perfectly acceptable when a client knows in advance they will be taking a new route. Often, however, in-car navigation systems would be most valuable precisely when the user **does not foresee** the need for assistance. For instance, imagine a driver taking an often traveled route back to work in Oakland from a sales destination downtown. Suppose further that this route is snarled by traffic congestion, but the driver is unaware of the upcoming situation. A system that can **predict** the driver's future destination and future route based on knowledge of the driver, the route taken thus far, and other features like time of day and day of the week, can provide a dramatic leap in the usefulness of computer assistance in vehicles.

In Figure 2, the model is applied to predicting a route in the Pittsburgh road network. The route originates near the Pittsburgh International Airport and terminates in a residential area near Carnegie Mellon University. Path costs were weighted using only simple features of the road network and not trained using empirical path preferences. However, already this model provides a reasonable probabilistic estimate of potential routes.



Figure 2. The probability distribution of different segments being contained in the route originating on the left of the map and terminating on the right. Greater intensity of red and greater width corresponds to higher probability.

**Integrating Probabilistic Components for Predictive Routing.** With a probabilistic model of paths for a user (or collection of users) in mind, we can go much further than simple margin based models as have been used previously [Taskar05, Ratliff06]. We can now invert the probabilistic model using Bayes' rule to answer questions like: "What is the probability of taking a particular path given the route taken so far and the destination?" and "What is the probability that the driver is trying to get to a particular location given the route taken so far and the current weather?"

While others have examined the problem of matching GPS traces to road segments of a road network [Letchner06] and detecting points of interest (destinations) from trace logs [Ashbrook02], a closer match to our proposed work in probabilistic modeling is the Microsoft PreDestination system that attempts to infer a driver's destination based on the route taken so far [Krumm06]. PreDestination divides a geographical area into a grid and makes a discrete probabilistic model of the probability of visiting those grid cells. It uses a driver's partial route to improve the estimate of reaching a potential grid cell destination: however, as it does not reason about the actual road network and conditions, it directly lowers the probability of cells that become further away. If the road out of your subdivision leads away from the grocery store, the grocery store will always become less probable as you drive away from your house.

The PreDestination approach is macroscopic, meaning that updates in destination belief can only be made when a vehicle has transitioned to a new grid cell. Our approach is instead microscopic, enabling updates in destination belief at each road segment traversed by the vehicle, which can lead to a more accurate prediction of destination earlier in the duration of the trip.

Perhaps most importantly, our approach dramatically extends this recent research [Krumm06] to be able to predict both destinations **and future routes** of a driver based on their past behavior and environmental conditions. This will allow us to reason about whether drivers will encounter adverse traffic or weather conditions that they may not be expecting. It will also lead to a system that can suggest improved routes or indicate to the user additional work that might easily be done along the way (e.g. the driver will be passing near the dry-cleaners where his clothes are currently waiting to be picked up.)

We will develop, validate and evaluate our purposeful planning approaches in two prototype products described below. They will be used also to acquire training data and to provide an initial dissemination of the techniques we develop.

**Web Interface and APIs for Route Planning.** With previously gathered datasets (20 Yellow Cab taxi drivers totaling over 100,000 miles, and 15 elder drivers for 3 months), we will design and train our machine learning algorithms for route prediction. We will first create a general model for routing, where it will treat all the provided data as belonging to a single user. Then we will modify our system to learn how to plan routes for a particular individual, treating the general model as a strong prior. With this system in place, we will provide access to it through a set of publicly available APIs and, in particular, a web interface, similar to that of Google Maps or MapQuest (see our mockup in Figure 3). The web interface will provide optimized routing under current or expected driving conditions for provided origin, destination and optional time/date of departure. It can customize the route, based on prior interaction with the website, and, if applicable, data collected from in-car GPS systems.



Figure 3. Sample web interface

**Mobile phone navigation system.** In addition to the web interface and APIs, we plan on developing a mobile phone navigation system. This system will leverage the customized planning system already described if a destination is provided, but can also predict the driver's destination and the route that the driver is planning to take, based on the path taken so far. Knowledge of the destination and route can be used to determine whether the driver is likely to be affected by unexpected traffic, weather or construction conditions. We plan to develop our navigation system on the Google Android cell phone platform, the most open platform for development and the expected dominant platform in 2 years. Using this platform, we can proactively deliver audio and/or



Figure 4. Android navigation application

textual prompts to the user about issues that may impact the driver. If drivers using the mobile phone navigation system are willing to share their driving data (either automatically using a wireless cellular network or manually uploading), then predicted routes for drivers could even take advantage of the experiences of other drivers. In the future, the ability to predict destination and route could also be used to slightly modify routes to help drivers complete necessary tasks (collected from personal information management tools like to-do lists and email) such as picking up dry-cleaning or buying milk.

Beyond the development of the web-based planning tool and the cell phone-based predictive navigation alerting tool, we plan to deploy both tools on a large scale to the Allegheny County driving population, to get large-scale, realistic usage and experience, with a concerted focus on supporting the Oakland-Downtown driving corridor. The evaluation will consist of surveys and interviews that measure accuracy and satisfaction of users, along with quantitative studies of the location and usage data collected.

## Outline of major tasks and rough timetable

As described in the previous section, there are 4 main tasks for this project: development of the probabilistic driving models for drivers in Allegheny County, development of a web-based route planning tool that leverages these models, development of a cell phone-based predictive system that leverages these models, and a significant deployment/evaluation of these systems.

**Table 1.** Rough timetable for the project (assuming start date of January 1, 2010)

<u> </u>	5	U	2 /	/	
Task		Jan-March	April-June	July-Sept	Oct-Dec
Probabilistic model development		Х	Х		
Web application development			Х	Х	
Mobile phone application development			Х	Х	
Deployment and evaluation				Х	Х

# **Planned Project Team**

The team consists of 2 researchers who have collaborated on the foundational work for the past 4 years.

Human-Computer Interaction	Anind Dey
Machine Learning	Drew Bagnell

# Potential for involvement of public or private partners

There is much interest in location-based services, particularly as a result of widespread adoption of GPSenabled cell phones. Our predictive navigation system is an example of such a service, and is of interest to multiple public and private partners. We already have a working relationship with the Yellow Cab taxi company in Allegheny County, and leveraged this relationship to collect our foundational dataset on which we based our initial technology development. Our technology and services will be of immediate interest to drivers for the Yellow Cab taxi company who frequently take fares along the Downtown-Oakland corridor. In addition, with the recent release of Google Navigation for the Android platform, and Google's interest in data-driven services and location-based services, Google is also a prime potential partner, and we have had initial conversations with Google Maps and Google Pittsburgh. We have also spoken with NavTeq, a recent acquisition of Nokia, which provides most of the online maps and maps data for North America. NavTeq has donated a short-term usage license to their data for Allegheny County, and we would be interested in extending this relationship.

In terms of public partners, we are looking to improve the data we have available for Allegheny County, such as the location of stop signs, left hand turn signals, number of lanes on a road, etc., which are not contained in the NavTeq data. We have spoken with a variety of people at the County and City levels, including the Assistant Director for Planning at the Allegheny County Economic Development, the

Allegheny County GIS Manager, and the PennDOT Portfolio Manager. We would be interested in partnering with any and all of these to improve the data on which our probabilistic approach is based.

#### **Cost estimate**

The estimated cost for this project is about \$100,000. This includes a month of time for both faculty investigators, 12 months (at a 50% rate) of a graduate student/postdoc, 1000 hours of hourly student programmer and usability/evaluation time, and funds for a non-research license to NavTeq's Allegheny County map data (if we are unable to partner with them).

### Potential sources of financial support other than Traffic21

We see a number of potential sources of funding beyond Traffic21. This includes the National Science Foundation CyberPhysical grants, Infrastructure grants, and CISE IIS (HCC and RI) grants. Our work is also relevant to the US DOT and PennDOT. We also expect that companies such as Google, mobile phone providers such as Nokia, and car companies that provide navigation support such as GM (e.g., OnStar) would have interest in funding our work.

### Laying the groundwork for future R&D funding

The work we are doing in probabilistic prediction and planning serves as a baseline for driver and transit support. Our future vision includes a much wider deployment and evaluation of our planning and prediction support for drivers. It also includes applying this technology to a number of related, but different projects. First, we will apply our predictive work to leverage information from multiple drivers, to not only use current location to perform traffic modeling, but to use the predictions of where those drivers are going and the routes they are taking to perform predictive and more accurate traffic modeling. Second, we will apply our work to predicting the paths that pedestrians will take for the purpose of controlling autonomous vehicles in pedestrian-heavy zones. Third, we will apply our work to transit riders, and predict ridership based on predicted routes and destinations of potential riders. Finally, we will look at environmental issues, and produce a series of applications that support green mobility. This includes producing routes that minimize fuel usage, and controlling HVAC and energy-heavy devices in the home based on predictions of drivers returning home. In particular, this last research avenue is of great interest to the Department of Energy.

#### References

- [Ashbrook02] Ashbrook, D. and Starner, T. (2002). Learning significant locations and predicting user movement with GPS. In Proceedings of the IEEE International Symposium on Wearable Computers.
- [CNN06]CNNMoney.com, Valdes-Dapena, P. Where those driving directions really come from. February 27, 2006. Available from: <u>http://money.cnn.com/2006/02/24/Autos/modern\_mapmakers/index.htm</u>
- [Hart68] Hart, P.E., Nilsson, N.J. and Raphael, B. (1968). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics SSC4* (2): 100–107.
- [Krumm06] Krumm, J. and Horvitz, J. (2006). Predestination: Inferring Destinations from Partial Trajectories. Proceedings of the Eighth International Conference on Ubiquitous Computing (UbiComp 2006): 243-260.
- [Letchner06] Letchner, J., Krumm, J. and Horvitz, E. (2006). Trip Router with Individualized Preferences (TRIP): Incorporating Personalization into Route Planning. *Proceedings of the Eighteenth Conference on Innovative Applications of Artificial Intelligence (IAAI-06)*.
- [Ratliff06] Ratliff, N., Bagnell, J. and Zinkevich, M. (2006). Maximum Margin Planning. *Proceedings of the International Conference on Machine Learning*: 729-736.
- [Taskar05] Taskar, B., Chatalbashev, V., Guestrin, C. and Koller, D. (2005). Learning structured prediction models: A large margin approach. *Proceedings of the Twenty-Second International Conference on Machine Learning*: 896-903.