# ONLINE ROUTE PREDICTION FOR AUTOMOTIVE APPLICATIONS

**Agnė Brilingaitė, Christian S. Jensen***

Department of Computer Science, Aalborg University, Denmark

* Fredrik Bajers Vej 7E, DK-9220 Aalborg Øst, Denmark, csj@cs.aau.dk, phone: +45 96358900

## ABSTRACT

An information and communication technology infrastructure is rapidly emerging that enables the delivery of location-based services to vast numbers of mobile users. Services will benefit from being aware of not only the user's location, but also the user's current destination and route towards the destination. This paper describes a component that enables the use of geo-context. Using GPS data, the component gathers a driver's routes and associates them with usage meta-data. Other services may then provide the component with a driver ID, the time of the day, and a location, in return obtaining the likely routes for the driver.

## KEYWORDS

Geographical context awareness, routes, destinations, travel patterns.

## INTRODUCTION

We continue to experience rapid advances in several key technology areas and in the diffusion of these technologies. In particular, these technologies include communication and geo-positioning technologies, as well as consumer electronics. An infrastructure is emerging that enables mobile users to subscribe to a range of advanced services. It becomes possible to deliver the right information at the right time to the service users.

Context-awareness then becomes a fundamental concept. If the services understand the context of the user, they can better deliver information that is relevant to the user, and its delivery will require less interaction with the user. Both are important benefits, one important reason being that the user is typically engaged in a primary activity, such as driving safely, that the mobile service should not distract the driver from.

Current location-based services [13] typically use the user's current location as the geo-context. We go further by focusing on the routes (and destinations) of the users as geo-context. People often use the same routes quite frequently and according to predictable patterns. For example, many people go to and from work at predictable times and use only a few routes for this.

When the routes of a user are known, a mobile service system can provide information related to the routes. If the system knows the route currently being followed by a user, it becomes able to provide information specific to this route. For example, the user will not be supplied by traffic information that does not affect the user's route. As another example, existing systems direct the users to the points of interest (e.g., gas stations) that are the closest to their current positions. In contrast, a system that knows a user's route may direct the user to the points of interest that is closest to the part of the route ahead of the user, thus minimizing the distance and/or detour needed to arrive at the point of interest [7]. As yet another example, a route-enabled system may alert the user about congestion and low-moving traffic ahead on the route.

We present a software component that is capable of continuously maintaining a prediction of a user's route and destination as the user drives. At the same time, the component receives position samples from a GPS receiver and accumulates route information, resulting in new routes being recorded in the system and in improved usage information for already known routes. Techniques already exist for the construction of routes from sequences of road segments by means of GPS data [2]. Thus, the details of route construction are not discussed in the paper.

The component can be made part of a larger system that uses knowledge about a user's route to provide a variety of improved, context-aware services. Examples include reservation of parking and provisioning of traffic information. Knowing the route of a user, such services can be made more relevant while reducing the user interaction needed. However, this paper focuses on the component as a stand-alone system, and it does not go into the more general use of route information in context-aware services.

Given the identity of a user and a time and location, the component predicts the user's destination and route to that destination based on prerecorded routes for the user and on information about the previous uses of these routes, i.e., their usage frequencies and temporal usage patterns. Some routes are used only during specific days of the week or specific times during the day.

As a user drives, the component maintains a ranked list of the user's possible routes. The component may use different options for updating this list. For example, it can recompute the list when the highest ranked route deviates and the current movement of the user no longer coincide. The paper offers examples how the component works based on a real road network and real GPS data.

The paper is structured as follows. In the next section, we describe the functionality, scenario, and settings. Then the section Management of Possible Routes covers route filtering. This is followed by the section Empirical Evaluation that addresses implementation and evaluation. The penultimate section covers related work, and the paper is ended by a summary.

## COMPONENT OVERVIEW

We first describe the stand-alone use of the component, then describe the technical setting within which the component works. Finally, we describe the component's internal functionality.

**Use Scenario** We assume a user who has a mobile device, e.g., a personal digital assistant or a navigation computer, with the route component installed. When the user starts the car, the component is activated, and it displays a list of the known, possible routes that start at the current destination.

The routes are ordered according to how likely they are to be the route that the user will be following. Initially, the possible routes are ordered based on the current time and on temporal usage patterns that describe when and with which frequencies the routes were used in the past. The route at the top of the list is the predicted route.

As the user starts driving, the system continually checks whether the user follows the predicted route. If the user deviates from the route, the list of possible routes is updated, and a new prediction is made. At any time, the user can select a destination in the system or select a route from the current list of possible routes. The system then updates its predictions based on this input.

**Technical Setting** First, all locations that a user visits (i.e., remains at for at least some specified duration of time) are termed destinations. A route then has a start and an end destination that are connected by so-called route elements, each of which is a road segment. Usage meta-data is also associated with each route.

A GPS receiver is used for identification of the road segments that make up a route. We thus assume that each mobile device has access to a GPS receiver. We also assume that a digital road network is available, and we map match the GPS positions to the digital road network.

For generality, we assume a client-server architecture, where mobile devices communicate with a server via some wireless network, e.g., GPRS. The assignment of functionality to the client and server sides is kept flexible so that the system can be adapted to different use situations. In one extreme, all functionality can be assigned to the clients. In particular, if the mobile device has ample processing power and memory, it can perform the map matching and route prediction. In this case, the server may simply need to occasionally synchronize with the client in order to use route information when providing the user with services. For devices with memory, processing, or power restrictions, the server may take over tasks.

**Internal Functionality** Figure 1 shows the main internal functions of the route component. Some functions are combined into groups (shaded regions), and some are kept separate. The first group of functionality covers the initial phase where the component starts, identifies the start destination, and produces a list of possible routes. The second group relates to the tracking of the user along a predicted route by filtering the routes. If the user deviates from the predicted route, the list of possible routes is updated, and a new prediction is made. The third group covers the construction of a new route.

When the user chooses a route from the list (the event *Input Route*), the system stops making predictions. When the user selects a destination (the event *Input Dst*), the system considers only those routes that end at this destination. If there is only one such a route, this becomes the predicted route. If there are several such routes, they are ordered and the highest ranked route is the one predicted.

## MANAGEMENT OF POSSIBLE ROUTES

**Ranking of Possible Routes** There are several ways of identifying possible routes and of ranking these. Adopting a naïve approach, all the routes that emanate from the start destination are possible, and the most frequently used one is the predicted route. For example, routes from home to work and back are the most frequent for many drivers.

Some routes have very distinctive temporal use patterns. To be better at predicting these, temporal usage patterns that capture the approximate time of the day, the day of the week, and the type of the day are associated with routes. These patterns also capture the usage frequencies.

For example, a “HOME–WORK” route may be used in the mornings on weekdays. With this temporal approach, the predicted route is the one that has a pattern that best fits the current time (time of the day, day name, and day type).

Some routes exhibit no clear usage patterns. For example, a user may visit family during some weekends, but far from every weekend. Such routes then are predicted only when initially predicted routes are eliminated as the user drives and deviates from these routes. The route to the family is one among several routes used during the weekends, with other routes relating to shopping, downtown visits, and visits to the seaside or countryside.

**Updating the List of Possible Routes** As mentioned earlier, the list of possible routes is updated as the user drives. Figure 1 covers the updating that occurs when a predicted route deviates from the actual route. If no specific data structure is used to store the routes, each pos-

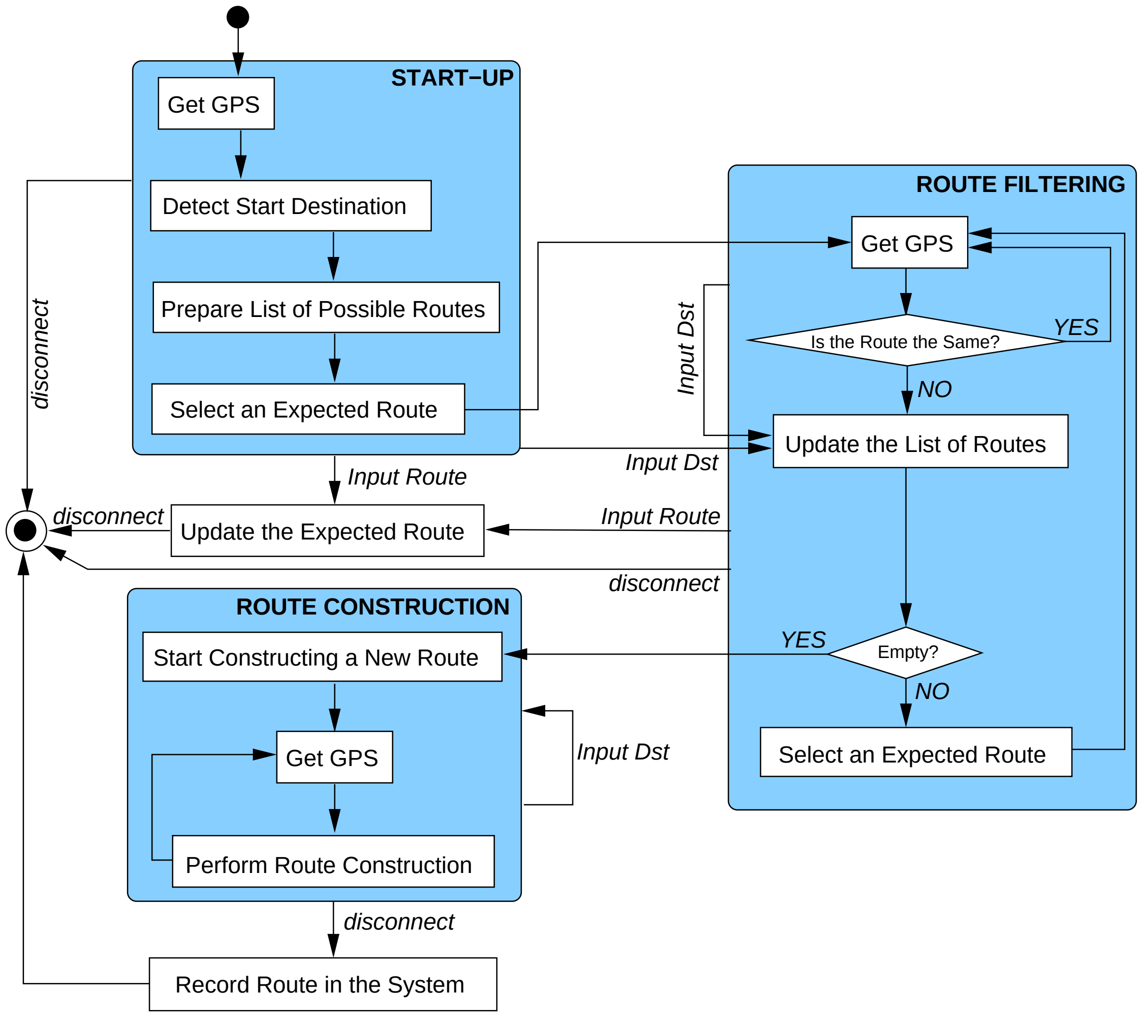


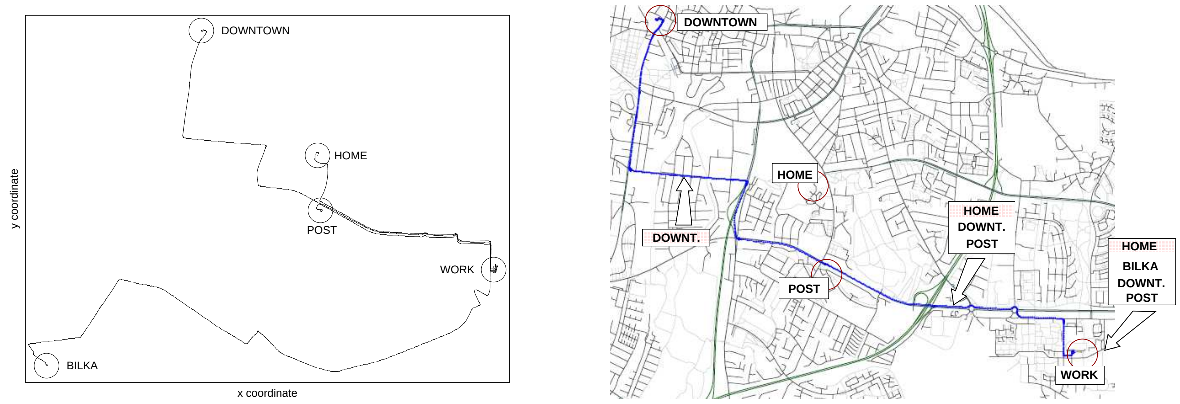

**Figure 1 – UML State Diagram—Internal Functionality of the Route Component**

sible route must be checked to determine whether it is still possible given the user’s most recent movement. We instead use an efficient data structure for maintaining the currently possible routes. This structure makes it efficient to remove routes that no longer belong to the collection of possible routes that coincide with the user’s movement so far.

## EMPIRICAL EVALUATION

To evaluate the component, it was implemented and used with the AKTA GPS data [6, 8], as well as with GPS data we collected ourselves, and the TOP10DK map of Denmark.

**Case Study** Figure 2 shows four routes that all start at the destination WORK and end at destinations HOME, POST (postal office), DOWNTOWN, and BILKA (a shopping center). As shown in Figure 2(a), the route to BILKA overlaps with the three other routes only at the beginning. However, the three other routes overlap for some time, until they start their own different sequences of route elements at an intersection. By turning left, the user reaches the close-by postal office; by turning right, the user reaches home; and by going straight, the user moves towards downtown.




(a) Four Routes—Four Destinations (b) Following One Route

**Figure 2 – Example of 4 Routes that Start at Destination WORK**

Figure 2(b) shows how the system reacts as the user starts driving from WORK. In particular, it displays the list of possible routes at each point along the actual route taken by the user. When the user starts driving, all four destinations are possible, and the route to home is the predicted one. Shortly thereafter, there are only three possible destinations, and when the user crosses the previously mentioned intersection, there is only one route, the one to downtown.

Figure 3 exhibits usage patterns (days of the week) for four routes from HOME to different end destinations. The first route is to WORK. It is almost always used on weekdays. The second

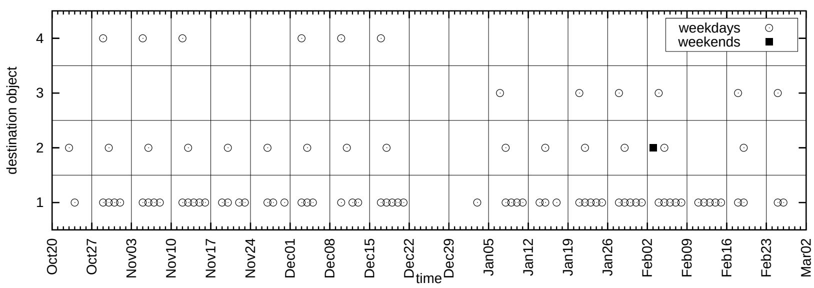



**Figure 3 – Time Patterns for Four End Destinations**

route is always used on Tuesdays. The third and fourth routes have similar use patterns (they are used on Mondays), but concern different destinations. At some point (at the end of the year), the user stopped going to the third destinations and instead started going to the fourth destination.

### Evaluation Using the AKTA Data Set

The naïve and temporal approaches to prediction were evaluated with the goal of understanding how the approaches work for a large data set. We selected 182 users that each traveled during a period of more than 80 days during the AKTA experiment and made more than 40 trips during those days. Using the first and last positions of the trips, destinations were identified by the automatic clustering of close positions. The resulting destinations were circular regions with radii of at least 150 meters, with the radii of a few destinations reaching more than four kilometers. This is due to the clustering procedure causing user destinations to be merged into larger destinations.

Temporal usage patterns were built for each pair of destinations connected by a trip. The number of possible end destinations reached from the same start destinations varied from 1 to 86 when considering only the guessable cases. If considering only those cases where the end destination was used in at least 10% of the trips from the particular start then the number of end destinations was between 0 and 10.

The naïve approach made correct predictions in 64% of all guessable cases, with 2% of the predictions having the same probability as the second candidate end destination. Thus, the correctness was at least 62%. When considering only the first half of the test period, these numbers for the approach are 66% and 63%, respectively. The correctness of the naïve approach depends on the usage frequencies of the predicted destinations. For example, half of the predictions were correct when the predicted destination had a frequency in the range $[0.45, 0.55)$, and the higher the frequencies of the predicted destinations, the higher the correctness.

The temporal approach considers not only the usage frequencies, but also the time of the day and the name and type of the day. This enables correct prediction of end destinations with low frequencies, in absolute terms and relative to other destinations. The approach utilizes several parameters during pattern construction. One parameter defines how much patterns for different day types are allowed to overlap. Another parameter defines how large the intervals that represent the starts of trips can be in a pattern. Yet another parameter is used to construct patterns of a particular length. The final main parameter controls whether patterns should be combined based on the distance between them. The larger these parameter values become, the fewer patterns there are; but the smaller these values are, the more "weak" candidate patterns with low frequency there are.

The temporal approach is better than the naïve approach at predicting low frequency destinations. For example, the naïve approach is unable to predict (correct) destinations with frequencies around 0.1, while the temporal approach is able to predict half of them. The average frequency of the predicted destination is 0.65 for the naïve approach and 0.59 for the temporal approach. The studies show that the correctness of the predictions for a user, stabilizes approximately one month after the first trip was recorded. The results show that it is possible to choose the parameters for the temporal approach so that the approach gives the best results when applied to the data.

## RELATED WORK

The work is related to several research areas, most notably tracking, location prediction, and travel surveying. A number of works have considered the tracking, or prediction, of the current and near-future positions of mobile users based on the most recently received movement information received for the users, e.g., position and velocity [3, 16]. In contrast, we continuously predict the route and end destination as a user moves.

Studies of the probabilistic modeling of user movement have also been reported. Ashbrook and Starner [1] discuss how GPS data can be used to predict the next location (destination) of a user. A Markov model is build for each location with transitions to all other locations. Liao et al. [12] build a probabilistic model for the transportation mode (on foot, by bus, by car) and the transitions of the user behavior. They use a hierarchical Markov model. Hariharan and Toyama [5] model location histories, i.e., location usages during intervals of time. They use location histories for Markovian and non-Markovian probabilistic models in order to predict future locations. We build temporal patterns for each start-end pair of destinations. Our iterative approach permits prediction update during a trip.

Kostov et al. [10] predict travel destinations using movement histories and time information. Entropy is used as an evaluation criterion. They use a hierarchical structure for time: The time of the day is divided into morning, afternoon, and evening, which are sub-divided further. They also distinguish between holidays and weekdays. The time intervals in our temporal patterns are constructed dynamically, and we distinguish in more detail among the days of the week when building and identifying patterns.

Karimi and Liu [9] predict the subsequent movement of a user based on unconditional probabilities assigned to road at intersections. In contrast, we consider the whole trip when predicting the next road choice. Laasonen et al. [11] identify personal important locations in cellular networks and predict abstract movement directions. We predict the end destination of a trip as a user moves.

Some studies also consider the identification of the purposes of trips. Assignment of the purpose of a trip to a destination has been studied [14, 15]. Schönfelder and Samaga [14] identify destinations using knowledge of public parking areas and shopping and leisure facilities. Purposes include "home," "work," and point of interests. A purpose was identified based on the time and on the distance from the last GPS point of the trip to points of interest. Wolf et al. [15] consider the replacement of traditional travel diaries by GPS data. Using GPS data and paper diaries they attempt to identify the purposes of trips. They used GPS data to manually identify destinations, which were street addresses. We do not rely on additional knowledge, and our goal is to predict the geographical destination rather than the purpose of a trip.

Doherty et al. [4] discuss the potential of GPS data for obtaining travel patterns of users. They require the users to confirm the trip purpose or the destination. We do not require any user input. Our system is functional without any user interaction.

## SUMMARY

The paper concerns the design, implementation, and empirical study of a route and destination prediction component. The paper covers the architecture and main functionality, it covers the approaches utilized for route and destination filtering, and it illustrates the use of the component by means of a real-world case study. Further, the paper reports on results of extensive empirical studies with the so-called naïve and temporal approaches to prediction.

**Acknowledgments** The second author is also an adjunct professor at Agder University College, Norway.

## REFERENCES


[1] Ashbrook, D., T. Starner (2002). Learning significant locations and predicting user movement with GPS. In *Proc. of the 6th International Symposium on Wearable Computers*, pp. 101–109.

[2] Brilingaitė, A., C. S. Jensen (2006). Enabling Routes of Road Network Constrained Movements as Mobile Service Context. *GeoInformatica*. To appear.

[3] Čivilis, A., C. S. Jensen, S. Pakalnis (2005). Techniques for Efficient Road-network-based Tracking of Moving Objects. *IEEE Trans. on Knowledge and Data Engineering* vol. 17, no. 5, pp. 698–712.

[4] Doherty, S. T., N. Noël, M.-L. Gosselin, C. Sirois, M. Ueno (1999 (2001)). Moving Beyond Observed Outcomes: Integraing Global Positioning Systems and Interactive Computer-Based Travel Bahavior Surveys. In *Conference of the Transportation Research Board Conference on Personal Travel: the Long and Short of It, Circular No. E-C026*, pp. 449–466.

[5] Hariharan, R., K. Toyama (2004). Project Lachesis: Parsing and Modeling Location Histories. In *Proc. of the Third International Conference on Geographic Information Science*, pp. 106–124.

[6] Homepage of the AKTA road pricing project (2006). www.akta-kbh.dk.

[7] Huang, X., C. S. Jensen (2004). In-Route Skyline Querying for Location-based Services. In *Proc. of the 4th International Workshop on Web and Wireless Geographical Information Systems*, *LNCS* vol. 3428, pp. 120–135.

[8] Jovicic, G., O. A. Nielsen (2003). The AKTA road pricing experiment in Copenhagen. In *10th International Conference on Travel Behaviour Research*. 33 pages. http://www.ivt.baug.ethz.ch/allgemein/pdf/nielsen.pdf.

[9] Karimi, H. A., X. Liu (2003). A Predictive Location Model for Location-based Services. In *Proc. of the 11th ACM International Symposium on Advances in Geographic Information Systems*, pp. 126–133.

[10] Kostov, V., J. Ozawa, M. Yoshioka, T. Kudoh (2005). Travel Destination Prediction Using Frequent Crossing Pattern from Driving History. In *Proc. of the 8th International IEEE Conference on Intelligent Transportation Systems*, pp. 970–977.

[11] Laasonen, K., M. Raento, H. Toivonen (2004). Adaptive on-device location recognition. In *Proc. of the 2nd International Conference on Pervasive Computing, Pervasive*, *LNCS* vol. 3001, pp. 287–304.

[12] Liao, L., D. Fox, H. A. Kautz (2004). Learning and Inferring Transportation Routines. In *Proc. of the 16th National Conference on Artificial Intelligence*, pp. 348–353.

[13] Schiller, J. H., A. Voisard, Eds. (2004). *Location-Based Services*. Morgan Kaufmann.

[14] Schönfelder, S., U. Samaga (2003). Where do you want to go today? - More observations on daily mobility. In *the 3rd Swiss Transport Research Conference*. 31 pages.

[15] Wolf, J., R. Guensler, W. Bachman (2001). Elimination of the Travel Diary: An Experiment to Derive Trip Purpose Grom GPS Travel Data. In *80th Annual Meeting*. Transportation Research Board. 22 pages, Paper. No. 01-3255.

[16] Wolfson, O., H. Yin (2003). Accuracy and Resource Concumption in Tracking and Location Prediction. In *Proc. of the 8th International Symposium on Advances in Spatial and Temporal Databases*, *LNCS* vol. 2750, pp. 325–343.