Prediction and Scheduling in Navigation Systems

Michael Minock
KTH: Royal Institute of Tech.
Stockholm, Sweden
minock@csc.kth.se

Johan Mollevik
Umeå University, Dept of CS
Umeå, Sweden
mollevik@cs.umu.se

ABSTRACT

This position paper makes a case for the need to *predict* pedestrian position and *schedule* communication acts in mobile navigation systems. In our work, carried out in the context of a voice guided city navigation system, we have found that improperly timed route instructions are a major cause of failure in guiding pedestrians in unknown environments. Furthermore, the need to communicate other information while guiding users on routes, as well as complications caused by network latencies, occurs often enough to require that we be able synchronize communication acts with user position as they follow a route. This has led us to focus our efforts on scheduling utterances to maximize route following success.

In this position paper we motivate this problem and present our initial approach and findings which should be of interest to others engaged in similar efforts in both the Geography and HCI communities.

Author Keywords

location-based systems; natural user interfaces; navigation systems; pedestrian interfaces; open street maps

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Algorithms, Design, Human Factors, Reliability

INTRODUCTION

The automated generation of route directions has been the subject of many recent academic studies (see, for example [10, 9, 4, 8, 11, 5, 2, 6]) and commercial projects (e.g. products by Garmin, TomTom, Google, Apple, etc.). While most focus has been dedicated to automobile drivers, there has also been an effort to provide route directions to pedestrians (e.g. Google and SIRI). The pedestrian case is particularly challenging because the location of the pedestrian is not just restricted to the road network and the pedestrian is able to quickly face different directions. In addition the scale of the pedestrian's world is much finer, thus requiring more detailed data representation. Finally the task is complicated by the fact that the pedestrian, for safety, should endeavor to keep their eyes and hands free – there is no room for a fixed dashboard screen to assist in presenting route directions. We take this last constraint at full force – following [1], in our prototype there is no map display; the only mode of presentation is text-to-speech instructions heard incrementally through the pedestrian's earpiece.

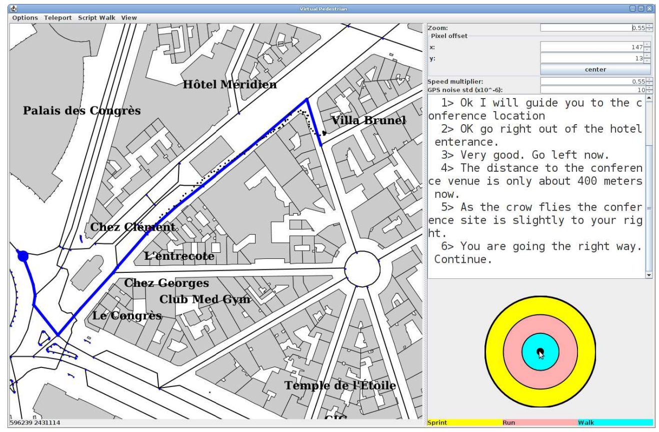

Figure 1. Guiding a pedestrian on a route near GeoHCI's venue.

Thus we focus on the problem of providing incremental spoken route directions to guide a pedestrian from their current position to a given destination. Such a problem yields a direct metric of evaluation: *what is the system's effectiveness, measure in time to destination or minimized deviation from route, in guiding pedestrians from a given initial position to a given destination position?*; The narrow position that we argue in this paper is that measures on this metric will be boosted considerably for systems that explicitly predict user positions and use scheduling to synchronize utterances with user position. We shall ultimately test this position by systematically comparing a *scheduling approach* with a *reactive approach*. The broader position argued for in this paper is that general user position prediction and scheduling of communication actions can be supported in modular components that can, without difficulty, be integrated into a variety of location-based navigation applications.

CONCEPTS

Consider Figure 1. Here we see a map, based on an export of OpenStreetMaps data[3], of a portion of Paris near to CHI's conference venue. We also observe a path that demarcates a route that the user is following (we shall follow the terminology of [10]). On the right-hand-side of Figure 1 we see a record of utterances that have been issued to the pedes-

GeoHCI Workshop at *CHI 2013*, April 2728, 2013, Paris, France.

trian to guide them along their path. Since our goal is to maximize the probability that the user follows the path, there is a mix of simultaneous *strategies* that are employed with little rest between utterances.

While the main strategy is to describe the turning actions (*turn sharp right now!*, *you will be turning right in about 20 seconds*), another strategy is to give positive feedback to encourage the user that they are pursuing the right path (e.g. *You are going the right way, continue walking.*). Another strategy is designed to inform the user that they are walking in the direction of their goal (e.g. *You are facing directly in the direction of your goal. It is approximately 400 meters away*). Yet another strategy provides descriptions of what the user should be seeing along the way (e.g. *you should see a large white building about 200 meters in the distance.*).

THE SCHEDULED APPROACH

Let us consider how one could support these simultaneous strategies and contrast two primary approaches: a *reactive approach* versus a *scheduled approach*. In the *reactive approach*, the system waits until the user arrives at certain points or, more generally, their trajectories meet certain conditions. At such points events are triggered which result in utterances being generated and voiced on the device. If this is implemented on the server side, then there will be some latency before the actual voicing of the utterance. Also if a particularly long utterance is being voiced, or if the user is moving more quickly than anticipated (e.g. on a bicycle), then the system may in fact miss presenting turn instructions in time. We have implemented a reactive approach earlier [7] and it often had such problems.

We contrast a reactive approach with a more sophisticated, *scheduled approach*. In such an approach, a schedule of future utterances is maintained. The most important utterances are associated with turning actions at decision points. Still, given that often the user will be traversing a path segment, other strategies also have room for their associated utterances to be scheduled. Utterances have start times, durations and pragmatic effects (e.g. enabling the user to correctly turn at a given branching point). The start times are projections into the future for when a given utterance will be issued. Once this time becomes equal to the current time, plus predicted latency, the call to voice the utterance is invoked. Obviously scheduled utterances may not overlap in time.

An interesting aspect of the scheduled approach is that within it one must represent a model of the user from which to generate predictions. This includes predicting the pace that the user will follow the route as well as the effect that utterances will have on their path. Such models can be more or less sophisticated. A simple model, that we term *inertial*, assumes that the user follows instructions perfectly and that their speed is constant (determined by sampling). That is the user continues in their given direction at their given pace, and responds perfectly to turn commands. Strictly speaking, such a model does not compel anything other than the most basic turning utterances. A slightly more complex model, where we assume that the user has a probability of making wrong turns (or failing to make turns), will explain the addition of extra utterances so long as these utterance's pragmatic effect is modeled as decreasing the probability that the user makes a wrong decision. There are many interesting user models to be developed around this problem, and much to be evaluated empirically with real pedestrians. One particularly interesting avenue of work will involve learning probabilistic models of the user from large samples of observed user data.

Other interesting issues are algorithmic and systems oriented. For example, in the most general case, we will wish to calculate a schedule that maximizes expected utility over a probabilistic user model. Note that the given user model may be considered orthogonal to the scheduling algorithmic, so long as it (the user model) is probabilistic. The computational complexity of the scheduling algorithm must be reduced to within bounds that allow real-time deliberation on modern hardware. In addition we will consider what parts of the calculation should occur on the mobile client and which on the server. Finally there are a whole host of issues around deciding when we should call for rescheduling of utterances.

INITIAL TECHNIQUES AND DEMONSTRATION

We have developed an Android-based platform for incrementally presenting spoken route directions to guide pedestrians to destinations. Our approach [7] makes heavy use of stored procedures and triggers in an underlying POSTGIS spatial database. In fact most of the 'intelligence' of our prototype resides in database stored procedures and tables. We have a base line reactive system as well as an initial scheduling approach. The initial scheduling approach uses an inertial user model for predictions. We are actively developing newer more sophisticated user models and we are also improving our scheduling and rescheduling algorithms.

We will be able to demonstrate our system live to interested GeoHCI participants in Paris. That we will guide interested persons on a tour around the region of the conference venue. In addition we will present a video demonstrating how we built the spatial database for Paris including the definition of tours and the authoring of various utterances.

CONCLUSIONS

We have presented here our ideas on the necessity to schedule utterances for users of navigation systems. Embedded within this requirement is developing user models that can be the basis of prediction. We suspect that this argument will also apply to more general multimodal interfaces as well. Finally we anticipate that the scheduling part of the algorithm will cleanly separate from the predictive user model part, and that alternative configurations of these components will suit different architectures, platforms and applications.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 270019 (SPACEBOOK project `www.spacebook-project.eu`) as well as a grant through the Kempe foundation (`www.kempe.com`).

REFERENCES

1. Bartie, P., and Mackaness, W. Development of a speech-based augmented reality system to support exploration of cityscape. *Transactions in GIS 10*, 1 (2006), 63–86.
2. Boye, J., Fredriksson, M., Götze, J., Gustafson, J., and Königsmann, J. Walk this way: Spatial grounding for city exploration. In *Proc. 4th international workshop on spoken dialogue systems, IWSDS'2012* (Paris, France, November 2012).
3. Coast, S. How OpenStreetMap is changing the world. In *proc. of W2GIS* (2011), 4.
4. Dale, R., Geldof, S., and Prost, J.-P. Using natural language generation in automatic route description. *Journal of Research and Practice in Information Technology 37*, 1 (2005).
5. Duckham, M., Winter, S., and Robinson, M. Including landmarks in routing instructions. *J. Locat. Based Serv. 4*, 1 (Mar. 2010), 28–52.
6. Janarthanam, S., Lemon, O., Liu, X., Bartie, P. J., Mackaness, W. A., Dalmas, T., and Goetze, J. Integrating location, visibility, and question-answering in a spoken dialogue system for pedestrian city exploration. In *SIGDIAL Conference* (2012), 134–136.
7. Minock, M., Mollevik, J., and Åsander, M. Towards an active database platform for guiding urban pedestrians. Tech. Rep. UMINF-12.18, Umeå University, 2012.
8. Miyazaki, Y., and Kamiya, T. Pedestrian navigation system for mobile phones using panoramic landscape images. In *SAINT* (2006), 102–108.
9. Nothegger, C., Winter, S., and Raubal, M. Computation of the salience of features. *Spatial Cognition and Computation 4* (2004), 113–136.
10. Richter, K.-F., and Klippel, A. A model for context-specific route directions. In *Spatial Cognition* (2004), 58–78.
11. Theune, M., Hofs, D., and Kessel, M. V. The virtual guide: A direction giving embodied conversational agent. In *Proc. of Interspeech 2007* (2007), 27–31.