Seems Familiar: An Algorithm For Inferring Spatial Familiarity Automatically

Norma Saiph Savage*, Wendy Chun*, Norma Elva Chavez**, Tobias Höllerer*

* Computer Science Department, University of California, Santa Barbara
** Facultad de Ingenieria, Departamento de Computacion, Universidad Nacional Autonoma de Mexico

Abstract. Our level of spatial familiarity with a particular place determines our navigation and interaction around that particular space. Despite its importance, most location based services rarely consider a user's spatial familiarity when delivering their services. The lack of integration of a user's spatial familiarity in location based services may be due to scarce work in automatic detection of a user's spatial familiarity.

In this work, following related studies in both psychology and geography, we propose a system that, given the user's current location, can automatically infer the person's level of spatial familiarity around the area he/she is currently in.

The system first maps the user's location to a specific place and obtains information about the place and the surrounding areas. The features considered for the classification are: The number of visits and time of last visit made to the location he/she is currently in, the type of place and the spatial layout associated with the place, the relationships between the user and the place, such as, authoritative constraints: a person under 21 cannot enter bars, as well as attributes related with the person's profession: a computer science major has visited numerous computer laboratories in the past, and therefore although they may currently be in a never before visited computer laboratory, to some degree, they are familiar with the environment. Finally the user's behavior is also considered for classification, in particular, a user's movement pattern, such as taking loops, shortcuts, or immediate reverse path (indicating that user took a wrong direction), and movement speed, such as a long pauses. The feature extraction is done on both a mobile device and external server. The classification is done with a Bayesian logic network in an external server. Our results suggest that it is possible to automatically infer a user's spatial familiarity given these features. We believe that the integration of our system to modern location based applications could improve the user's experience. For instance, map visualizations and automated trip generation applications could be dynamically changed depending on the user's level of spatial familiarity.

Keywords. Spatial familiarity, User Models, Inference

1. Introduction

The level of spatial familiarity we present with our surroundings is a key factor of how we navigate and interact with the environment. Yet most location based services (LBS) rarely consider this variable when delivering their services, because the work, which has considered this variable, has rather focused more on analyzing how a person's level of familiarity with their surroundings affects the performance of the person's tasks. For example, in Lovelace, Hegarty, et al., 1999, the authors proved that the type of route a user prefers to take highly depends on his/her familiarity with the area. Additionally the work done by Powell, S.B., M.A. Geyer, D. Gallagher, and M.P. Paulus, 2004, demonstrated that the degree of familiarity with ones surroundings defines the degree of exploration, i.e., in a familiar environment, that the individual presents a greater initiative to visit novel places in comparison to when he/she is in an unfamiliar environment.

We believe that using these results, we can implement very useful LBS applications in which, using known values of a certain area or a specific place based on the individual's level of familiarity, we can present different routes as well as map visualizations. Moreover, knowing whether a user is familiar or unfamiliar around the surrounding areas could also improve the generation of automated trips.

The underlying problem is that despite its importance, not much work has been done in automatic inferring system to find whether a person is familiar or unfamiliar with the environment he/she is currently in; advances may be lacking due to the fact that the concept of familiarity is evidently a much more complex concept than many other concepts that also qualify knowledge. The work of Nathan Gale, Reginald G Golledge, William C Halperin, Helen Couclelis (1990), intended to be the first approach in measuring a person's spatial familiarity. The authors devised two forms for qualifying these values, one was through self-evaluation and the other was through the analysis of a person's actions and behaviors.

Clearly, self evaluation would resolve the problem of knowing a user's current spatial familiarity, but constantly querying a person about how familiar he/she feels with each new place he/she visits could impose a cognitive load and could also be very cumbersome for the user, and it is therefore not a viable solution.

Recent work has taken the approach of analyzing a person's behavior to infer his/her degree of familiarity. The work done by Meneses, Filipe; Moreira, Adriano, 2006, is an example of this; the problem their work presented is that, despite the diverse behaviors that have been linked to spatial familiarity such as exploration and routes selected to traverse the space, the authors only considered the number of visits to a location for a measure of spatial familiarity. In this work, we propose a novel algorithm which analyzes much more features for automatically inferring a user's Spatial Familiarity with his/her current environment. In the following section we explain our algorithm in greater detail.

2. Algorithm Overview

Our algorithm is currently designed to return familiarity scores to places the user visits. For each user a user model is created and it holds information about the places the person has visited in the past, such as the computed familiarity score and other relevant place information. The user model also stores a user's personal information. All of these data is obtained from the user’s social network, in particular we utilize Facebook and foursquare. It was decided to obtain the information from social media sites, to limit the user’s cognitive workload. The only information the user is requestd to provide are her social media credentials and to list 3 places she can affirm to feel familiar with. These places are utilized for ground truth data. Our system adds these places into the user model, and when the system detects that the user is visiting these places, the system analyzes the user’s behavior in these places as well as the relationship between the user and the place, to understand what makes a familiar place familiar.

When the user visits a new place, our system first obtains information about the place. Our system utilizes information available on the internet around the place for the data gathering. In specific, foursquare is utilized, that is, from foursquare, the nature or category of the place is obtained. This category is then analyzed against the categories of the places the user has visited. We count how many places in the user’s history match the category of the current place the he/she is visiting. From this number we obtain its logarithm, the result provides a metric indicating how familiar the person feels interacting with the space. This score is stored and is later utilized for calculating the overall familiarity score. The next step in our algorithm is to analyze if the place's category matches the category of places that are known to have generic layouts, such as airports. We manually mapped each foursquare category to a score indicating whether or not the layout of the places belonging to that category is known to present the same spatial navigation space. This score is stored and is later utilized for calculating the overall familiarity score. The next step in our algorithm is to infer if the place is related in some way to the user's profession or to the user's hobbies. Psychological studies have shown,(see for example J M Cortes, a Greve, a B Barrett, M C W van Rossum, 2010) that people express familiarity sentiments when the place they are in is related to their works or to some of their leisure activities and that is why our algorithm integrates this metric for measuring spatial familiarity. To obtain information about the user's job and hobbies, we crawl the user's Facebook profile page. (to use our application, the user must provide their Facebook and foursquare credentials and give the application access to their data.) From the user's Facebook profile, we obtain the person's education information, e.g., studied computer engineering in the National Autonomous University of Mexico, jobs held, e.g., worked as software engineer for Google, and the information in activities and interests, e.g.,playing beach volleyball. From the words in these different entries, a document holding all the words is created. Next, the foursquare tags and comments that are associated to the place the user is currently in are retrieved. For each word associated with the place, its log frequency weight[28] with respect to the document that holds the user's work and hobbies information is calculated. The log frequency weight of a word w with respect to a document d can be defined as a function f(w):

$$f(w) = \begin{cases} 1 + \log\ wf_{w,d} \\ 0 \end{cases}$$

Where $wf_{w,d}$ represents the number of times the word w occurs in the document. All the log frequency scores are then summed together. This value represents a user's familiarity with a place, given the user's professional and personal life. This score is stored and it will later be utilized for calculating the overall familiarity score

The last variable our algorithm analyzes is the user's movements. From the user's GPS traces obtained from their mobile device, it analyzes to find whether the user's trajectory presented loops or if the user took an immediate reverse path after advancing forward a bit. We also analyze if to reach their destination the user took shortcuts. Studies in Psychology and geography have shown that the use of shortcuts expresses a degree of spatial familiarity (see for example Eckhardt, A. and Vojt_ 2009, Kirill V. Istomin and Mark J. Dwyer, 2009). Based on the user's movements, a score between one and zero is assigned. The movement analysis is based on the approach taken by Y. Ishikawa, Y. Machida, H. Kitagawa (2006) and N.S. Savage, et.al., (2010). This score is stored and it will later be utilized for calculating the overall familiarity score.

At this points our algorithm has a series of familiarity scores, and these scores will be utilized to obtain an overall familiarity score for the place the user is in. For this, we compare each particular score, with scores of places that the user has explicitly said they were familiar with (The scores of the places the user has expressed to be familiar with, are calculated when the user visits the location). The distance between each of the scores with respect to the scores of a labeled familiar place is computed; all distances are then added and divided over one. The familiarity metric formula is as follows:

$$
\begin{aligned}
&SumOfAllDistances \\
&= distance(FamiliarityScoreAOfLabeledPlace, FamiliarityScoreAOfCurrentPlace) \\
&+ distance(FamiliarityScoreBOfLabeledPlace, FamiliarityScoreBOfCurrentPlace) \\
&+ \cdots distance(FamiliarityScoreNOfLabeledPlace, FamiliarityScoreNOfCurrentPlace)
\end{aligned}
$$

$$OverallFamiliarityScore = \frac{1}{SumOfAllDistances + 1}$$

This familiarity metric is based on the first law of geography that states: "Everything is related to everything else, but near things are more related than distant things." The greater the distance that exists between the familiarity score of a place the user stated she is familiar with, and the familiarity score of the place the user is currently in, the smaller the familiarity metric. Our formula has one added to it, that is, when the distance is zero, our algorithm returns a familiarity metric of one.

For modeling the familiarity metric values over time a Bayesian belief network is utilized.

3. Conclusions

In this work, we presented a novel algorithm for automatically inferring the spatial familiarity of a user to a given place and surrounding area. Our method utilized information from social networks and a user's mobile device.

The next stage of our project is to do a user study to evaluate our algorithm. Our current study shows that, a user's spatial familiarity can be automatically inferred.

Measuring a user's spatial familiarity is very useful for various location based services, and we believe our algorithm should be integrated into these applications so that a user experience can be enhanced. For example, in places the user is spatially familiar with, the system can provide less navigation aid, while when the user is in an unfamiliar space, the system can provide additional help. Knowing the user's spatial familiarity can aid systems in deciding what notifications and assistances to provide to the person without overwhelming them. And despite its preliminary character, our system presents a first step to automatically inferring a user's spatial familiarity.

Acknowledgments

This work was partially supported by a UC MEXUS-CONACYT fellowship, as well as NSF grants IIS-0747520 and IIS-1058132.

References

Kristin L. Lovelace and Mary Hegarty and Daniel R. Montello, 1999, Elements of Good Route Directions in Familiar and Unfamiliar Environments, Spatial Information Theory

Powell, S.B., M.A. Geyer, D. Gallagher, and M.P. Paulus. 2004. The balance between approach and avoidance behaviors in a novel object exploration paradigm in mice. Behav. Brain Res. 152:341–34

Nathan Gale, Reginald G Golledge, William C Halperin, Helen Couclelis , 1990, Exploring Spatial Familiarity. in The Professional Geographer

Tobler W., (1970) "A computer movie simulating urban growth in the Detroit region". Economic Geography, 46(2): 234-240.

Meneses, Filipe; Moreira, Adriano; 2006 , "Using GSM CellID Positioning for Place Discovering," Pervasive Health Conference and Workshops, vol., no., pp.1-8, Nov. 29 2006-Dec. 1 2006

Y. Ishikawa, Y. Machida, H. Kitagawa (2006) ,A Dynamic Mobility Histogram Construction Method Based on Markov Chains, 18th International Conference on Scientific and Statistical Database Management (SSDBM'06) p. 359-368

Norma Saiph Savage, Shoji Nishimura, Norma Elva Chavez, and Xifeng Yan. 2010. Frequent trajectory mining on GPS data. In Proceedings of the 3rd International Workshop on Location and the Web (LocWeb '10). ACM, New York, NY, USA, , Article 3 , 4 pages

Eckhardt, A. and Vojt_a_s, P.: Combining Various Methods of Automated User Decision and Preferences Modelling. Lecture Notes in Computer Science, Volume 5861/2009, pages 172-181, 2009.

Kirill V. Istomin and Mark J. Dwyer, 2009, Finding the Way: A Critical Discussion of Anthropological Theories of Human Spatial Orientation with Reference to Reindeer Herders of Northeastern Europe and Western Siberia, Current Anthropology, Vol. 50, No. 1 (February 2009), pp. 29-49

J M Cortes, a Greve, a B Barrett, M C W van Rossum (2010), , Dynamics and robustness of familiarity memory.Neural computation 22 (2) p. 448-66 Bertin J (1967) Semiology of graphics: diagrams, networks, maps. University of Wisconsin Press, 1983 (first published in French in 1967, translated to English by Berg WJ in 1983)