Bootstrapping the Location-enhanced Word Wide Web

Bill N. Schilit[1], Anthony LaMarca[1], David McDonald[4], Jason Tabert[3], Eithon Cadag[3], Gaetano Borriello[1,2], William G. Griswold[3], Jason Hong[5]

[1]Intel Research Seattle, 1100 NE 45th Street, Seattle, WA 98105
[2]Dept. of Computer Science and Engineering, University of Washington, Seattle, WA 98195
[3]Dept. of Computer Science and Engineering, University of California at San Diego, La Jolla, CA 92093
[4]Information School, University of Washington, Seattle, WA 98195
[5]Computer Science Division, University of California at Berkeley, Berkeley, CA 94720

ABSTRACT

Our challenge to the research community is to make location-enhanced web services valuable and readily accessible to a very large number of people in daily, real world, situations. We envisage a global scale, multi-organization and interdisciplinary initiative, *Place Lab*, that will bootstrap the broad adoption of the location-enhanced Web. Our research collective is developing an open software base (providing low-cost private positioning technology) and fostering the formation of user and developer communities. Through individual Place Labs initially seeded on the campuses of universities, colleges, and research organizations this initiative will be a vehicle for research, instruction, collaboration and application sharing. This paper describes some of our first steps towards meeting this challenge.

Keywords

Location-aware; context-aware; ubiquitous; positioning systems; WiFi; GPS; web services; wireless hotspots; wardriving.

INTRODUCTION

Location-aware computing has been "in the lab" for the last decade where applications, frameworks, technology infrastructure and much more has been extensively explored. So far location-based services have not made an impact in the mobile computing world (except perhaps with E911 services). The reason is the ubiquitous computing dilemma: how to bootstrap a concept that requires both infrastructure investment and the "killer" (or at least a valuable) application. Without the application people won't invest in infrastructure and without the infrastructure the open-market for iterating towards valuable applications and their business models doesn't exist.

Place Lab is a community-based effort to break this deadlock and make location-aware computing a reality on a mass scale. We see at least three major barriers that must be overcome to realize this vision: low-cost, highly convenient position-sensing technology; making users comfortable with respect to their location privacy; and having existing web content easily customized to geographic locations.

Figure 1: WiFi density in urban centers is such that multiple access points are within range of many locations. Each AP beacons a unique identifier that, along with a mapping database, can be used to lookup a course grain position. In this image each dot is an estimate of the position of a WiFi AP in downtown Seattle mapped in a single "wardrive."

Our current approach exploits the proliferation of wireless networking hotspots that can provide positioning comparable to GPS in urban settings and also function indoors where GPS does not. A downloaded and continually updated distributed contributor database of all the WiFi access points in the world will allow clients to compute their own positions and divulge their location information only when they want to. Services accessed through a web browser will provide users rich information and services associated with their location.

The Place Lab "research challenge" provides an endeavor where the lessons from the field of location aware computing can be applied. This knowledge includes the idea of location-enhanced web browsers, proposed in 1995 [2], and the extensive contributions around WiFi for location [1]. We believe, however, that in order to take location from the laboratory to the real world there remain significant research challenges. In this paper we first present a scenario, describe three challenges, and conclude with a preview of a Place Lab demonstration occurring at Ubicomp 2003.

USAGE SCENARIO

A Place Lab user subscribes to databases, potentially from multiple providers, that the client WiFi Positioning algorithms use to convert an access point BSSID (plus signal strengths) into a geographic position. We expect these databases would be updated once a week or so and might cover large geographic regions such as North America, Europe or Asia. Over time we see this collection of WiFi Positioning databases growing to include every access point in the world (later we describe some ideas on how to bootstrap and maintain the databases). Given such a collection of databases, whenever the client receives BSSID beacons they are able to calculate position without additional network communication. This client-based calculation of position-without-communication is a fundamental principal of the privacy mechanism proposed for Place Lab.

On visiting a location-enhanced web service, the user is able to trade privacy of their location for utility of the web service. We imagine a Place Lab component, the Place Bar, which integrates WiFi Positioning into the user's web browser and allows users to flexibly send location information at various fidelities enhanced sites. For example, the user might choose to reveal only one of these about their location: country; state (prefecture, canton, province, etc.); city; neighborhood; postal code; street; street address; and longitude/latitude.

HARD PROBLEMS

We have identified at least three hard problems that stand in the way of realizing Place Lab:

1. How to bootstrap and manage a worldwide hotspot database for positioning?

2. What is the trust model at the client, what is being revealed, and how can we avoid the "big brother" hot button?

3. How to associate any page on the web with a place in the real world where it might be useful? How can multiple pages appropriate for a location be organized for easy browsing?

These problems, and probably many more, must be addressed by the research community as this challenge moves forward. In the following sections we describe potential solution directions.

How to Bootstrap a Global WiFi Positioning Database?

The first technique to bootstrap a WiFi Positioning database is to generate war-driving data for a region, such as the UCSD campus and town of La Jolla. The idea is to create a rough, incomplete map of the hotspots in an area. With this database, notebook computer and PDA users without GPS can start contributing more information into the database. For example, assume a user goes to a Starbuck's and receives beacons from three APs but only two are in the database. The third AP can then be added to the database with some high confidence that it is near the

location of the other two APs. Data can also be added when an unknown AP is detected temporarily between two known APs. This collection of techniques for refining the details of the WiFi Positioning database as a side effect of people using their mobile computers is the second, geographic statistical technique.

Clearly, the data being collected by the geographic statistical technique would be much more useful if it was sent back into the infrastructure and then redistributed to all users as part of the WiFi Positioning database. The third technique is to employ a distributed contributor update mechanism for the WiFi Positioning database similar to the one made famous by the CDDB service:

The WiFi Positioning database could aggregate and statistically process AP sightings, and even use the distributed contributor model to improve the precision of the data over time. In some situations, users might be presented with the current location information that is being sent off to the location-enhanced web service. If users notice an error in the location, or the database just holds the city and not the street, the user could enter the corrected or more precise location information that would eventually be added to the database. Of course, users should not be able to corrupt the database. Statistical methods coupled with authoritative sources of hotspot location can be used to ensure high-quality.

What is the Trust Model & What is Being Revealed?

Whenever a location system is developed we can expect to hear shouts of "big brother!" Some of the news headlines that came out of the Active Badge location systems include: "big brother pinned to your chest," "Orwellian dream come true, a badge that pinpoints you," "badges monitor staff."

The privacy problem is due in part to the choices we present people: either opt-in or opt-out with no levels in between. When opting-in the systems we design generally send location to a central server, that we expect users to trust. Most users do not trust centralized location tracking servers run by the government, large corporations, or even your University's IT staff. As an example you can look at the debate over E-911 in congress.

For Place Lab the questions "when I'm using this what am I revealing?" and "when I'm *not* using this what am I revealing?" are make-or-break questions for adoption. Our approach is two fold: (1) client-only position calculation; and (2) multi-fidelity location revelation.

Client-only position calculation is the antithesis of the "big brother" location server: all computation of a device's location occurs at the trusted client. GPS is a good example of this model. In the case of Place Lab, the inputs to the computation are AP beacons received at the client and a cached copy of a database that allows mapping the WiFi beacons (possibly with signal strength data) to locations. At this basic level of WiFi Positioning, if a client does not use
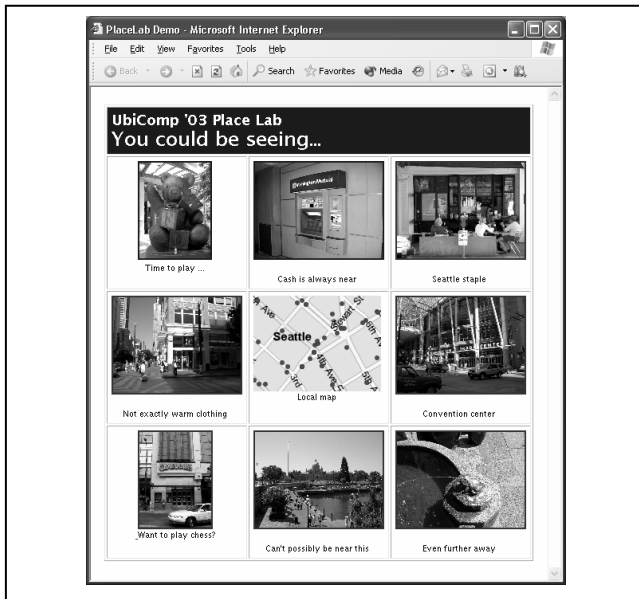
Figure 2: The main page of the Place-Enhanced Conference Guide presents images of interesting "sights" from around the conference venue (conference Hotel). The user's location is detected through WiFi hotspots that have been previously mapped. The content (images, factoids, opinions, and links) are both manually created, culled from the Web prior to the conference, categorized, geo-coded, and placed in an install package. When a particular sight is selected, more detailed information is displayed. The entire web site runs without network connectivity and uses beacons from the last seen WiFi hotspot to approximate location

the APs for communication, then a totally private positioning system is possible.

Generally, we think people would want to interact with some location-enhanced service or fetch some location-specific content so communication would be a normal part of the usage model. However, it is worth noting that the private disconnected mode is worth exploring. First, there are times when you may be near APs that are available for positioning but not available for communication (they might be private, belong to another provider, or maybe you didn't pay your T-Mobile bill last month). In these cases you might interact with offline web content in an "occasionally connected computing" (OCC) model supported by .NET and other frameworks. For example, if the Zagat restaurant guide was an OCC location-enhanced site, you could make the content available offline and use it with live location information. In this way, you could get information about nearby restaurants without actually revealing location data to the Zagat server.

How to Associate Web Pages to Places?

Another of the challenges in the place-enhanced Web is content discovery: "how do you find information associated with a location?" The most obvious approach is to ask content-providers to annotate their pages with location information. However, how is this "geocoding" structured? Are pages to be tied to specific coordinates? How big is the region around a coordinate for which the page is still relevant? To further complicate matters these regions are unlikely to be simple rectangles and will undoubtedly overlap with each other.

There are two main approaches to dealing with this problem: asking content providers (and third parties) to code their pages with location information; or deriving the locations associated with a page through observation of users' browsing habits. In the first case, we may not get many associations at all since we have put an extra burden on content providers. In the second case, we have to determine how privacy preserving aggregation techniques can be used to collaboratively associate pages with locations. An important issue is where to store and compute these associations.

Even when we have these associations in place, we must still tackle the problem of how to present this information to the user. What happens when a user asks for information associated with a place? What will they see in the browser? The associated pages will have to be indexed in some way. This may be by some of the hierarchies we discussed above but they could also be organized by what we know about the user's current context and what they are likely to be most interested in. This could be based on past experience (implying that we may employ some machine learning techniques) or be specific to more explicit information such as calendar and preference data.

UBICOMP 2003 DEMO

At UbiComp 2003 we are demonstrating a proof-of-concept system to launch our community development effort. We have developed a stand-alone system that conference participants can download and install onto their laptops that will give them a location-aware conference guide for the neighborhood that surrounds the Ubicomp '03

In our demo, users will interact with the conference guide via a standard web browser accessing HTML pages. The map view on each page will place the user on a map of downtown Seattle (or a detailed map of the conference Hotel). The page will also present images of nearby locales. The users can drill down from the basic view to find interesting images, facts and opinions.

One of our concerns designing the conference guide was that the location algorithms we are using provide rough grain information. Although we expect that in time other researchers will apply better algorithms to improve this aspect of Place Lab, we knew that it was possible that position reports could be off by a city block or more! Our first interface had a text-based style and included specific descriptions of computed position. We decided to generalize the interface with imagery, including a map

containing of few blocks, in order to avoid confusion if the positioning broke down.

CONCLUSION

A growing multi-organization group of researchers is developing the concepts, open code base, and collaborations that comprise Place Lab. We plan on seeding several partner universities with the necessary elements to develop Place Lab enabled applications and expect a variety of classes from different departments to start the development of relevant and valuable location-aware applications. Our near term objective is to create a way to share of applications across all campuses. Our long term goal is to break the cycle that is preventing location-aware usage models from developing on a large scale.

REFERENCES

1. Paramvir Bahl, Anand Balachandran, Allen K.L. Miu, W. Russell, Geoffrey M. Voelker and Yi-Min Wang, "PAWNs: Satisfying the Need for Ubiquitous Connectivity and Location Services," IEEE Personal Communications Magazine (PCS), Vol. 9, No. 1 (February 2002).

2. Geoffrey M. Voelker and Brian N. Bershad. Mobisaic: An Information System for a Mobile Wireless Computing Environment. In Proceedings of the Workshop on Mobile Computing Systems and Applications, pp. 185-190, December, 1994.