Nomatic: Location By, For, and Of Crowds

Donald J. Patterson, Xianghua Ding, and Nicholas Noack

University Of California at Irvine, USA
{djp3, dingx, nnoack}@ics.uci.edu

Abstract. In this paper we present a social and technical architecture which will enable the study of localization from the perspective of crowds. Our research agenda is to leverage new computing opportunities that arise when *many* people are simultaneously localizing themselves. By aggregating this and other types of context information we intend to develop a statistically powerful data set that can be used by urban planners, users and their software. This paper presents an end-to-end strategy, motivated with preliminary user studies, for lowering the social and technical barriers to sharing context information. The primary technology through which we motivate participation is an intelligent context-aware instant messaging client called Nomatic*Gaim. We investigate social barriers to participation with a small informal user study evaluating automatic privacy mechanisms which give people control over their context disclosure. We then analyze some preliminary data from an early deployment. Finally we show how leveraging these mass-collaborations could help to improve Nomatic*Gaim by allowing it to infer position to place mappings.

1 Introduction

A great deal of progress has been made in the pervasive computing community on the problem of localizing individuals. There are outdoor systems [1], indoor systems [2], and hybrid indoor-outdoor systems [3, 4, 5, 6]. There are infrastructure-based solutions [7] and wearable/mobile-based solutions [8]. There are IR [9], acoustic [10, 11], laser [12], single-point [13] and sensor fusion techniques [14, 15]. Hightower presents an excellent survey [16] with other examples and categorizations.

With such a wide array of techniques for empowering a user to digitally leverage their location, it is now justified to begin studying new computing opportunities that are enabled when *large numbers* of people begin using these technologies. Urban planners have implicitly begun studying these issues through the analysis of where cell-phones are carried and used [17, 18]. We view our work as complementary to theirs: crowd context can certainly be used for urban analysis. However, we wish to expand the types of analyses that can be done by providing statistically powerful data sets which can still be used by urban planners, but also by the crowds who generate the data, and the software that they

M. Hazas, J. Krumm, and T. Strang (Eds.): LoCA 2006, LNCS 3987, pp. 186–203, 2006.

use. A key component of our vision is to enable aggregations over place (as opposed to position [19]), activities and social situations.

To be specific, as a first order goal, we wish to develop methods of being able to answer location-aggregation questions such as "Where are UCI students right now?" One answer to such a question is to show positions on a map, but our different and novel approach is to answer in the form of a probability distribution over *places*: "90% of UCI students are *at home*." Different questions might be answered with distributions that can be interpreted as: "This weekend, 50% of the assisted living center residents went to a *store*", or "Right now, 10% of the sales force are in *developing nations*". Our future goals include being able to answer activity- and social status-aggregation questions that naturally spring to mind as well.

The risk versus benefit trade-off for a user who reveals context information changes in two ways when it is aggregated. The first change is that the privacy dynamic is altered. An observer trying to understand information about crowds does not need access to knowledge about a particular person. As a result, some of the risks of losing plausible deniability [20], and of being surveilled [21] are lowered. They certainly don't go away (see section 3 for example!) because in order to aggregate someone must have access to individual context information, but anonymity becomes more practical and sufficient statistics can be maintained without maintaining data on individual behavior. The second change is that motivation to reveal context information changes as well. Since no personal analysis of context is available to the user once their data is aggregated, the benefit of revealing that information is small. Therefore new ways of encouraging participation in the community from which aggregations are made must be developed.

The rest of our paper structures the social and technical architecture which we are developing to achieve our goal. The centerpiece of the architecture is our context-aware instant messaging client, Nomatic*Gaim, described in section 2. We argue that this application needs context-awareness immediately and, as a result, is sufficiently motivating for people to reveal context information. Nonetheless, this revelation is a *social* barrier that we do not take lightly. As a result, in section 3 we present the results of a small informal user study that we have conducted to help shape our design of privacy controls to further facilitate participation. Creating a well-designed IM client is a *technical* barrier that we investigate in a small deployment described in section 4. Encouraging results from these studies provides a suggestion that these barriers can be crossed, so in section 5 we show how remote data collection conducted by Nomatic*Gaim will support solving the position to place problem, and thus the aggregate query challenge posed above. Additionally, this particular solution can be used to further lower social and technical barriers to the use of Nomatic*Gaim and other location-aware systems.

2 Nomatic*Gaim

The prefix tag, "Nomatic*" refers to our system for collecting descriptions of context correlated with position. Nomatic has the particular characteristics of leveraging mass collaboration to get statistically powerful amounts of data which can then be used as data for data mining algorithms to solve challenging cultural and social context representation problems.

2.1 Our Use of Context

We use “context” as short hand for three types of data: a user’s current place, a user’s current activity, and a user’s current social situation. While our approach could include other aspects of context as well, these three have immediate promise with regard to Instant Messaging (IM) usage.

– **Place:** Place is a way of describing a position in a semantic way. It is an inherently ambiguous and subjective label that depends on who is labeling a place, why they are labeling it and who they think is going to see the label. Unlike place, position is an exact location, that, while possibly difficult to gather, clearly describes where something is. It is usually described in a coordinate system such as latitude and longitude, but may also be described in other ways based on use and technology. Regardless of representation, position unambiguously maps where locations are (e.g., 31N -117W, 3 miles north of Exit 14 on highway I-5, etc.). Position to place is not a one-to-one mapping. Not only can a position have many place names, but a single place name may map to many positions (e.g., home, work, a Yoshinoya restaurant, an IKEA store). An excellent discussion of the issues surrounding position and place can be found in [19].
– **Activity:** Activity is also a subjective description of what a person is currently doing. Much of the same ambiguity that surrounds place also surrounds activity. A person may be walking, talking on a cell-phone, having a conversation, laughing, exercising, looking for the subway, all at the same time.
– **Social Situation:** This is the way in which a person would describe their current activity as it relates to other people. It includes situations such as “being alone”, “being in a crowd”, “in a meeting”, “on a date”, etc. Again, this is a highly subjective way of describing part of a person’s context.

In fact, these three elements of context are closely tied together. To say that you are at a theater likely means that you are also watching a movie, and in a crowd of strangers. To say that you are flying to Sydney also implies that you are on a plane. So we do not suggest that these three types of context are independent, but rather they are different lenses through which we observe situations.

When necessary, some of the ambiguity of context can be eliminated by conforming to an ontology, hierarchical or otherwise. GIS systems, such as the U.S. Census bureau’s [22], frequently impose such a structure on the description of places, so that each position has exactly one well characterized label, perhaps “light industrial”, “park”, or “shopping mall.” Much work in the pervasive and ubiquitous computing community on activity recognition takes this tactic as well. Many algorithms are made tractable by the exclusive and unambiguous description of activities in progress (e.g., [23, 24]).

Nomatic, through its mass-collaborative nature, rejects the imposition of a strict ontology and supports the emergent creation of communal understandings of context that will arise through shared use. We allow users to arbitrarily tag locations in the style of de.li.cious [25], Flickr [26], and Etsy [27] so that Nomatic can grow and change with community use.

2.2 Instant Messaging Transformed

The first step in motivating crowds toward a mass-collaboration of context information disclosure is to provide an incentive for individuals to reveal it. We believe that a transformed IM client is just such an incentive.

The Human-Computer Interaction and Computer Supported Cooperative Work literature has thoroughly documented the value and use of "awareness technologies" to support and improve distributed group work [28, 29, 30, 31, 32, 33]. Ethnographic studies of IM have shown that the awareness associated with the online/offline status line of an IM client has substantial value in the maintenance of dyadic human relationships [34, 35, 36].

However, the computing context in which IM has been used has been rapidly changing. Until recently most users of IM were primarily using desktop computers. Computing has changed substantially in recent years such that laptops sales are outpacing desktop computer sales [37] and IM clients are now available on mobile phones. This means that the context provided by a label that says "online" vs. "offline" or "available" vs. "not available" is no longer sufficient to achieve the positive externalities mentioned in Nardi's work [36]. Some studies indicate that 13% of all IM dialog is simply related to negotiating availability [38]. Increasingly, individuals are always "online". The difference is that now a person might be "online" and only available for some kinds of IM messages. She may be driving, giving a presentation, using the same hardware to make a phone call, or incur fees to receive IM messages. "Online", no longer means that a user can gracefully accept all interruptions.

To adapt to this new reality, we have developed a context-aware, open-source[1], cross-platform[2], and cross-protocol[3] instant messaging client called Nomatic*Gaim. It is the merger of two existing projects, Gaim [39] and Place Lab [6], with the Nomatic mass-collaboration context collection system. Unlike previous IM clients, this one reports more nuanced context information in the status line (see figure 1). When a user sets their context information, their position is simultaneously queried from the Place Lab system. These two pieces of information are then alternately reported on the awareness status line where the "available" label was previously. (Figures 2–3 shows other clients displaying Nomatic*Gaim status).

[1] Nomatic*Gaim is licensed under a combination of GPL, and other redistributable licenses.

[2] It compiles and runs on Windows, Linux and Mac OS platforms.

[3] MSN Messenger, Yahoo! Messenger, AIM/AOL/iChat, Jabber/Google, and others.

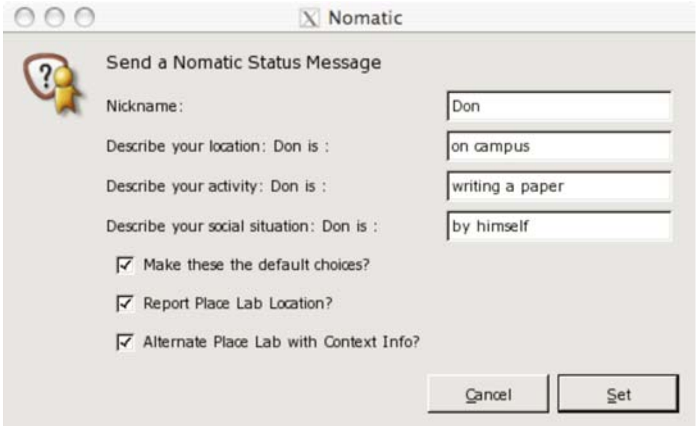

Fig. 1. Nomatic*Gaim's context entry dialog for manual context entry

Fig. 2. Apple iChat client displaying status created by Nomatic*Gaim

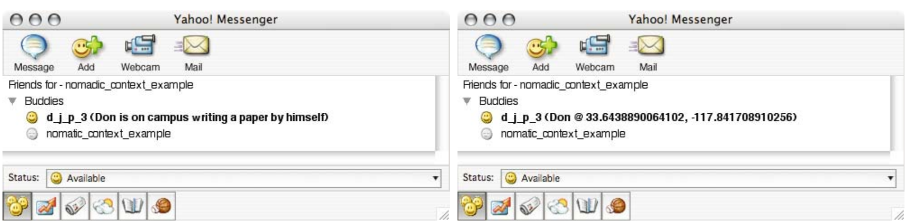

Fig. 3. Yahoo! Messenger client displaying status created by Nomatic*Gaim

At the same time as the current context information is reported to a user's buddies, the Nomatic*Gaim client also anonymizes the (context,position) pair and sends it to the Nomatic Mass-Collaboration Database (NMCDB) where a database entry is made matching the position with the manual context information. In this way the user is provided immediate benefit while a database of position to place name mappings and, more generally, position to context mappings is being developed.

As exciting as this new way of using IM is, there are some clear concerns about privacy that need to be addressed. Lack of appropriate controls over context disclosure may make users uncomfortable about adopting Nomatic*Gaim's awareness model. In order to understand user attitudes and to develop models for automatically managing privacy, we conducted a small Wizard-Of-Oz pilot study to direct our efforts.

3 Privacy Pilot Study

We made the hypothesis that a user would be motivated to use Nomatic*Gaim's context disclosure system if they had appropriate control over the method in which the information was released. To evaluate this, we conducted a small pilot study prototyping various IM privacy configuration management strategies. For this study, we focused on what benefits users can achieve from location disclosure through IM, what concerns this entails, and what factors affect their disclosure attitudes. Although this study is small and informal, we hope it will iteratively guide more formal user studies in the future.

Previous work has exposed some of the underlying privacy issues associated with location disclosure. The private nature of location information and users' unwillingness to disclose it has caused some research efforts to fail [9, 40, 41]. Some have found that automated location disclosure is not very well accepted by users, and concluded that automatic features lacked value [42]. Others have determined that people first choose whether or not to reveal location, and if so, then specify their location in the most useful terms for a given relationship [43]. Still others have supported this idea that relationship dominates disclosure decisions [44].

Part of why we believe we may be able to succeed where others have not is because we suspect that the underlying technology has a significant effect on how people choose to reveal context information. Most previous work has focused on SMS[4] communication, but IM represents a different mode of communication, and affords different information practices (i.e. different degrees of synchronicity, central vs. ambient attention demands, "pulling" vs. "pushing" modes of operation, etc.). The employment of different communication media for location disclosure may change the balance between benefits and costs, affecting a user's sense of control, and thus their attitudes.

The issue of privacy in our IM client is partially mitigated by controls built into the IM protocols themselves which allow users to control access to their online/offline status, and therefore Nomatic*Gaim status, to people whom they

[4] Short Message Service: text messages sent to cell-phones.

Fig. 4. Three customized paper prototypes of potential future Nomatic*Gaim interfaces used in the pilot study. The left interface is manual. The right interface is completely automated. The center interface gives the most control to the user in disclosing their current place.

have previously approved. The discussion of context disclosure is therefore made within the boundaries of preexisting privacy mechanisms. (see [45] for more details on IM privacy by protocol).

3.1 Methodology

Our study consisted of two parts, a paper-prototyping/Wizard-Of-Oz study and a scenario-based questionnaire evaluating potential location-aware IM usage models for a laptop computer. The paper prototypes demonstrate possible functionality that differs from the existing Nomatic*Gaim client.

The study was conducted on the UCI campus in November 2005. Restricting ourselves to students within our research group who were not working on Nomatic, we screened potential participants to ensure they used IM daily and were mobile laptop users. We recruited five students (3 male-2 female, 4 graduate-1 undergraduate) who received a short application form that collected information to customize the subsequent scenarios.

During each one-hour session, we introduced three paper prototypes, presented several scenarios that required users to interact with the prototypes, and then conducted a semi-structured interview with the participants to understand their experiences.

Privacy Interface Options. The three paper prototypes were designed to give the users a spectrum of automation over location disclosure. In the first one, the system doesn't do any automated disclosure, but users can manually enter and disclose their current place to their entire buddy list. In the second one, the system automatically detects a variety of place names that describe the user's position. The user is then able to drag place names over groups of individuals to reveal their place. The third one automatically chooses and reveals a place name to all the people in the buddy list. As with the scenarios, the interfaces were customized for each participant. An example is shown in figure 4.

3.2 Discussion

Our screening interviews reflected a wide variety of practices and environments involved with IM usage. One unexpected situation involved IM usage during church services. Opinions of showing location information also varied during the study. One participant began their session by saying that location disclosure is not something they wanted and didn't have any value to them. However, she later expressed several times that she had changed her mind, and agreed that this feature could be very convenient and useful.

Benefits. Through the study, the following four benefits emerged as recurring themes:

- **Facilitating Activity Coordination:** Nomatic*Gaim has the potential to let others know where you are and vice versa. This appeared to be useful for pre-defined, but flexible activity coordination (e.g, impromptu office meetings, coffee shop meetings and project meetings before a class). Our interviews demonstrated that these would be the highest value scenarios for the participants as they were useful, convenient and efficient.
- **To Facilitate Socially Appropriate IM Interactions:** Most participants agreed that location information is a good indicator of a user's social context. Providing location information through IM can inform and signal others how to interact more appropriately. Seeing an IM buddy at a meeting location was interpreted as a signal that their buddy was waiting for them. Additionally, "Meeting Room", "Church", and "Classroom" locations all indicate strong expectations over interruptibility. It also signals different expectations for different people. For example, one participant emphasized that the "Church" label means he doesn't mind people from the same church contacting him (before the service starts), but not others. The observation was also made that for people who are familiar with each other's schedule, place information reveals what types of activities they are engaged in and provides cues for interruptibility.
- **Emotional Benefits:** One participant mentioned that his parents call him all the time, just wanting to know where he is. Another participant stated that "There are probably 2-3 people I'd like to disclose [my location] to all the time", and another, "I guess I also care what my parents know, because they are most likely to give me crap over the telephone, or to worry (i.e. I'd like them to know I am home after a plane flight)" These statements suggest that location information revealed through IM could enhance social connections. Unlike phone calls and SMS, IM costs less in terms of time and attention and may achieve better benefits as a result. Two participants mentioned that this feature could save time by reducing the frequency of check-up calls from parents. As one participant suggested, it provides "background" awareness of her remote friends.

- **Increased Opportunities for Socializing:** Location information can lead to opportunistic social interactions. One participant mentioned that when she sees her friend is in New York, she will start a conversation by “Hey, what are you doing in NY?” Another participant suggested that “Home” means he is available for his friends. Yet another participant mentioned that if she sees her friends at “Home”, she will “feel more comfortable to reach them, without any worries”. In our coffee shop scenarios, participants expressed their willingness to disclose their location information if they are alone, “so [friends] nearby can join me”.

Information Leakage. Besides benefits, the scenarios and interview questions also expose some tensions and concerns involved. Some participants were concerned that location disclosure might cause information leakage over scope and time, and lead to undesirable consequences.

Our participants demonstrated an unwillingness to disclose some location information to conceptual groups of people in their buddy list. As one participant put it, “a couple [people] I’d actually want to hide information from, I’m protective of my face time.” In a home scenario, one participant said “office mates don’t need to know [where I am] for sure”. Some locations are especially sensitive to some people such as “bars” to parents, or “home” to boss, although as a result of current IM practice, participants also mentioned that these people were currently not in their buddy list.

One participant was afraid of information leakage over time, either choosing to disclose or choosing not to disclose location can be informative “in ways you don’t like.” The example given is “parents asking ‘why I didn’t display ‘Home’ all weekend?’ Even if it doesn’t show where I actually am.” Another participant also pointed out choosing not to disclose location may cause social pressure. A third mentioned that people might learn her rhythms, and could predict her locations in an undesired way. System security was also raised as a concern.

High Level of Acceptance. Our pilot study showed a high level of acceptance of location disclosure through IM and the corresponding automatic features. All participants answered “yes” when asked whether they would use this feature if it were available. At the same time, two participants also mentioned that they would adopt it only when it became very popular suggesting that achieving critical mass will be an important social barrier to adoption. Nothing about the proposed system requires multiple users to participate.

The high level of acceptance exceeded our original expectations, and is contrasted with findings from previous location disclosure systems. We speculate several reasons why this might be:

- **More Benefits:** IM makes location exchange very easy, which is significant for activity coordination. Secondly, although existing IM clients support quick and informal conversations, they suffer from a lack of social context which location disclosure helps to mitigate.
- **Lower Attention Cost:** One major feature of IM is its presence awareness. It works like an ambient display [46], operating in the background, utilizing

peripheral attention, and supports easy movement between background and foreground operation. This feature significantly lowers the attention cost on the information receivers, and makes it less annoying and interruptive.
- **Sufficient Level of Control:** In contrast with cell-phones and PDAs, laptops are less intimate. People typically do not bring them everywhere they go. This eliminates some of the threats to a person's privacy. Secondly, the buddy list provides a natural and controllable boundary in terms of the scope of information leakage. Finally, the awareness metaphor associated with IM status disclosure is simple to conceptualize.

3.3 Future Design Directions

- **Temporality:** Several participants distinguished between regular and one time events, and suggested different ways of automating each. For scheduled and regular activities, automatic features have the potential to reduce tedious manual specification. Also, with regular activities, participants seem to have less privacy concerns. In contrast, for one time events, manual disclosure is acceptable or even preferred.
- **Spatiality:** In some places, such as work, our participants showed their willingness to use automatic features to reveal location information. It had high value, and little privacy risk. In contrast, "home", or "coffee shop" were more private.
- **People Aspect:** People want to disclose their location to groups of people in different ways. Furthermore people categorize their acquaintances differently. This suggests a flexible buddy organizing mechanism is an important requirement for automatic location disclosure design.
- **Activity Aspect:** Our study showed that for the same location, different social situations altered attitudes towards location disclosure. For example, while waiting at a coffee shop, almost everybody indicated they would "definitely" choose to disclose location. However, after they start socializing, nobody wanted to disclose location anymore. When the socializing is over, if they remained alone, they all chose to disclose location again. This suggests the need to provide a very easy mechanism to quickly start and stop location disclosure.
- **Information Aspect:** People are less concerned about disclosing more generalized location information to their buddy list. For most participants, generalized information doesn't convey too much information, and yet is useful enough for relevant people. Detailed information such as room numbers were only appropriate for certain people. This suggests using generalized information as default settings.
- **Easy Control:** Our participants made it clear that control had to be "dead easy" – "No more than one click to override."

3.4 Summary of Pilot Study

The size and informality of our study makes it impossible to make conclusive statements about privacy in IM. We did observe generally positive attitudes towards location disclosure from our well-informed and computer literate user sample. They also emphasized that automatic privacy mechanisms need to be flexible in regard to recognizing what times and places were valid to make default disclosures and to provide very clear feedback and control about what information was being revealed. We also got hints that made us suspect that critical mass, rather than privacy will be the determining factor in wide adoption of this feature.

4 Preliminary Nomatic*Gaim Deployment

In parallel with our work on privacy automation we tested an initial version of our system to evaluate technical barriers that would impact the effectiveness and system design of Nomatic*Gaim. We used an instrumented version of our client running on Windows and Macintosh laptops that collected statistics about how our users used physical space and IM. We had one professor and two graduate students use our client over the course of one week each. At the end of the usage period we collected and analyzed the resulting logs.

4.1 Collected Statistics

The first set of statistics that we collected related to protocol usage and supported our decision to use a multi-protocol approach. Figure 5 shows the number of buddies on each persons IM list separated according to protocol. These numbers are in line with similar research into IM usage that determined that an average buddy list has 22 people [35]. Figure 6 shows a similar analysis based on number of messages per protocol.

The second set of statistics that we collected was related to Wi-Fi Access Point (AP) localization. This is the underlying technology present in Place Lab which in turn is used by Nomatic*Gaim to determine a user's position. This is

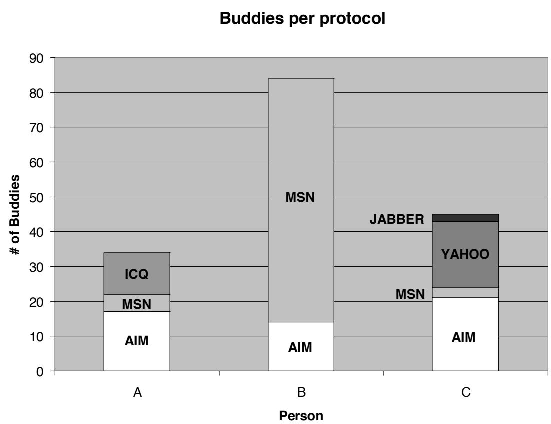

Fig. 5. Number of buddies per IM protocol in pilot deployment

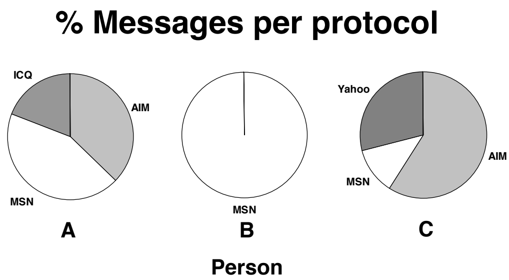

Fig. 6. Number of IM messages per protocol in pilot deployment

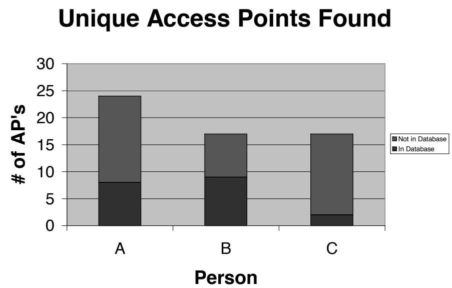

Fig. 7. Unique access points found by each user. Light gray reflects APs that were not in the database.

a beacon-based technology that compares the currently observed APs to a local database of AP locations. By triangulating a position among the observed APs, a user can passively ascertain his location with moderate accuracy.

We looked at how many unique APs our users encountered and what percentage of them were not represented in the Place Lab data base (see figure 7). Over the course of the week there was no period of time in which a laptop was used outside the range of an AP. This was despite the fact that one of the users in our study made a cross-country plane trip during the study period.

4.2 Discussion

– **AP Coverage.** The number of APs in the database was surprisingly low. There were always APs visible, but their location was frequently not known. Whenever this was the case, Place Lab was unable to find a latitude/longitude for the user. However, the set of AP signatures did suggest a particular location, just one that was unable to be placed in a coordinate system.
– **Accuracy.** The literature cites resolution accuracy that is very high under ideal situations, sometimes under 10m [6, 47]. Although we didn't formally analyze accuracy, our anecdotal experience suggested that we achieved approximately 30m accuracy in practice.
– **Mobility.** Consistent with the results of a study into the mobility of CMU students [48], we found that most laptop usage was in a few regular

repeated areas. Only during the cross-country travel episode did we see many locations per user.

4.3 Future Directions

These results suggest a few important features for the Nomatic system. First, bootstrapping and auto-extension of APs must be easy and integral to the system. A system like Microsoft Live Local [49] that accepts unknown APs and attempts to localize them based on co-location with known APs will be essential for the success of this system.

Additionally it would be convenient to have an interface that allowed the user to act as a manual GPS. Such an interface would allow them to pick a position on a map that is their actual position and then update the relevant databases to reflect the APs in sight.

Finally the mapping of position to place should not fail just because absolute position is not known. A graceful segue between known unique positions and absolute positions is critical while AP localization is in the bootstrapping stage.

5 Closing the Loop: Leveraging Aggregates

So far we have discussed a preliminary system, based on an IM client that will collect and reveal context information about a user, in a manner with which they are comfortable. We have argued that IM is a compelling application that provides immediate gratification to the user. It is less clear how a user would benefit from providing their context information to the NMCDB and why they wouldn't forgo that aspect of the system while they displayed the same information to their buddies.

Motivating the mass-collaboration aspect comes from the powerful ability for a Nomatic*Gaim client to use statistical inference from the data that has been submitted to the NMCDB. This inference can determine the best place label for the given position, L^*, based on the data collected so far at all nearby positions, x, and the currently observed location, z:

$$L^* = \operatorname*{argmax}_{L \in Places} \int_x P(L \,|\, Data(x)) P(x \,|\, z)$$

Since $P(x \mid z)$ is essentially a neighbor function, we can generalize position to be either a latitude and longitude pair, or a set of APs, and thus support bootstrapping. In the latitude/longitude case, the notion of a neighborhood is physical. In the AP case the notion of a neighborhood is "all locations that have a non-empty intersection of APs". We can assume that the greater the overlap the closer the locations.

We can further differentiate between the data that the user has personally collected and data that the rest of the community has collected and assume the two are independent:

$$L^* = \operatorname*{argmax}_{L \in Places} \int_x P(L \,|\, Data_{user}(x), Data_{\neg user}(x)) P(x \,|\, z) \tag{1}$$

$$L^* = \operatorname*{argmax}_{L \in Places} \int_x P(L \,|\, Data_{user}(x)) P(L \,|\, Data_{\neg user}(x)) P(x \,|\, z) \qquad (2)$$

We now consider the case of a user who has never used Nomatic*Gaim before. This user arrives at a new location and opens their laptop. Place Lab localizes the user's position. Nomatic*Gaim gets the position and sends it to the NMCDB. The NMCDB responds with the optimal place label for the given location, L^*, calculated according to equation 2. Since $Data_{user}(x) = \emptyset$, $P(L \,|\, Data_{user}(x))$ is uniform and uninformative.

After a user has entered information into the database on their own, their particular labelings of space are non-empty and begin to contribute toward the understanding of place for their IM client.

We can now incorporate the design guidelines that we have uncovered to create a better user experience. Rather than having the NMCDB respond with, L^*, it can respond with the list of highest rated place labels, $(L^* = L_0, L_1, L_2...)$ and the associated probabilities, $(P(L_0 \,|\, \hat{z}), P(L_1 \,|\, \hat{z}), P(L_2 \,|\, \hat{z})...)$. Nomatic*Gaim can now take one of four options under the assumption that $\alpha > \beta > \gamma$:

- If $P(L_0 \,|\, \hat{z}) > \alpha$ then Nomatic*Gaim will automatically set the place context for the user
- If $P(L_0 \,|\, \hat{z}) > \beta$ then Nomatic*Gaim will automatically set the place context for the user and ask for the user to confirm the setting.
- If $P(L_0 \,|\, \hat{z}) > \gamma$ then Nomatic*Gaim will open the context dialog with the list of locations, $L_0...L_n$, ordered according to their associated probabilities in a drop down box for user selection.
- Otherwise, Nomatic*Gaim will open the dialog box for the user and allow him to manually set the current location place.

The final user experience is compelling. If a user is willing to collaborate with NMCDB, when they arrive at a brand-new location and open their laptop, Nomatic*Gaim will take one of the proposed U/I options. If this new location has a common name that all previous people have used to label that location, the user's place (not position) status will be set for them automatically. They will never have to touch their IM client and a semantic understanding of their current location will be displayed. If the user doesn't like the common place labeling, and they correct it manually, the next time they arrive at the same location, their customized place label will automatically be used. The user never again has to touch their status when in that location. If for reasons of privacy or ambiguity a location has multiple place names, the user's current place will be confirmed before it is set for the user. The sensitivity to the automatic processes can be set by altering α,β, and γ.

6 Conclusions

In this paper we have presented a social and technical architecture for collecting large amounts of context information from users in real-time.

Through the use of several small user studies we have identified social and technical barriers to the adoption of our context-aware instant messaging client, Nomatic*Gaim. Our privacy study was surprisingly positive given related work in the field. We attribute this apparent willingness of people to participate in context disclosure to the unique features of IM that are different from SMS messaging. Our early deployment study supported our client design decisions of being cross-protocol and identified the need for techniques that work in the absence of absolute knowledge of position.

While previous work and technological trends suggest that the context awareness provided by Nomatic*Gaim has immediate benefits for a user and their buddies. It is crucial to the overall vision of Nomatic that users also reveal their context information to the NMCDB. To this end Nomatic*Gaim also leverages data submitted to the NMCDB to improve it's U/I, as an example of statistical inference over the data from crowds, but also to encourage continued mass-collaboration from the individual user.

Ultimately by leveraging this architecture, we are able to provide information that will assist urban planners, people and software to answer novel new queries that aggregate over the semantic contexts of crowds of people. This information can be used by people for aggregate social analysis, or by machines to create more social interfaces.

Acknowledgments

Thanks to all the folks who helped us out with this paper: Kah Liu, Matt Nguyen, Mike Moore, Grace Chen, Amanda Williams, Paul Dourish.

References

1. Hatch, R., Sharpe, T., Galyean, P.: Starfire: A global high accuracy differential gps system. Technical report, NavCom Technologies (2006)
2. Bahl, P., Padmanabhan, V.N.: RADAR: An In-Building RF-Based User Location and Tracking System. In: INFOCOM (2). (2000) 775–784
3. Griswold, W.G., Boyer, R., Brown, S.W., Truong, T.M.: A Component Architecture for an Extensible, Highly Integrated Context-Aware Computing Infrastructure. In: Proc. of the 25th International Conf. on Software Engineering, IEEE Computer Society (2003) 363–372
4. Griswold, W.G., Shanahan, P., Brown, S.W., Boyer, R., Ratto, M., Shapiro, R.B., Truong, T.M.: ActiveCampus - Experiments in Community-Oriented Ubiquitous Computing. Technical Report CS2003-0750, UC San Diego (2003)
5. Schilit, B., LaMarca, A., Borriello, G., Griswold, W., McDonald, D., Lazowska, E., Balachandran, A., Hong, J., Iverson, V.: Challenge: Ubiquitous Location-Aware Computing and the Place Lab Initiative. In: Proc. of WMASH 2003: The First ACM International Workshop on Wireless Mobile Applications and Services on WLAN. (2003)
6. LaMarca, A., Chawathe, Y., Consolvo, S., Hightower, J., Smith, I.E., Scott, J., Sohn, T., Howard, J., Hughes, J., Potter, F., Tabert, J., Powledge, P., Borriello, G., Schilit, B.N.: Place lab: Device positioning using radio beacons in the wild. [50] 116–133

7. Krumm, J., Williams, L., Smith, G.: SmartMoveX on a Graph – An Inexpensive Active Badge Tracker. In Borriello, G., Holmquist, L.E., eds.: Proc. of UBICOMP 2002. Volume LNCS 2498., Springer-Verlag (2002) 299–307
8. Krumm, J., Cermak, G., Horvitz, E.: RightSPOT: A Novel Sense of Location for Smart Personal Object. In Dey, A., Schmidt, A., McCarthy, J.F., eds.: Proc. of UBICOMP 2003. Volume LNCS 2864., Springer-Verlag (2003)
9. Want, R., Hopper, A., Falcão, V., Gibbons, J.: The active badge location system. ACM Transactions on Information Systems **10**(1) (1992) 91–102
10. Lopes, C.V., Haghighat, A., Mandal, A., Givargis, T., Baldi, P.: Localization of off-the-shelf mobile devices using audible sound: Architectures, protocols and performance assessment. Mobile Computing and Communications Review **to appear** (2006)
11. Scott, J., Dragovic, B.: Audio location: Accurate low-cost location sensing. [50] 1–18
12. Fox, D., Burgard, W., Thrun, S.: Markov localization for mobile robots in dynamic environments. Journal of Artificial Intelligence Research **11** (1999)
13. Wilson, D.H., Atkeson, C.G.: Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors. [50] 62–79
14. Hightower, J., Brumitt, B., Borriello, G.: The Location Stack: A Layered Model for Location in Ubiquitous Computing. In: Proc. of the Fourth IEEE Workshop on Mobile Computing Systems and Applications, IEEE Computer Society (2002)
15. Liao, L., Fox, D., Hightower, J., Kautz, H., Schulz, D.: Voronoi Tracking: Location Estimation using Sparse and Noisy Sensor Data. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE/RSJ (2003) 723–728
16. Hightower, J., Borriello, G.: A Survey and Taxonomy of Location Systems for Ubiquitous Computing (2001)
17. Polak, E., Society, W.: Amsterdam realtime (2002)
18. Ratti, C., Pulselli, R.M., Williams, S., Frenchman, D.: Mobile landscape: using location data from cell-phones for urban analysis. Environment and Planning B **to appear** (2006)
19. Hightower, J.: From position to place. In: Proceedings of The 2003 Workshop on Location-Aware Computing. (2003) 10–12 part of the 2003 Ubiquitous Computing Conference.
20. Smith, I.E., Consolvo, S., LaMarca, A., Hightower, J., Scott, J., Sohn, T., Hughes, J., Iachello, G., Abowd, G.D.: Social disclosure of place: From location technology to communication practices. [50] 134–151
21. Mann, S.: ”sousveillance”: inverse surveillance in multimedia imaging. In: MULTIMEDIA ’04: Proceedings of the 12th annual ACM international conference on Multimedia, New York, NY, USA, ACM Press (2004) 620–627
22. Bureau, U.C.: Census 2000 TIGER/Line Data. http://www.esri.com/data/download/census2000-tigerline/ (2000)
23. Patterson, D.J., Fox, D., Kautz, H.A., Philipose, M.: Fine-grained activity recognition by aggregating abstract object usage. [51] 44–51
24. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In Ferscha, A., Mattern, F., eds.: Pervasive. Volume 3001 of Lecture Notes in Com puter Science., Springer (2004) 1–17
25. (Del.icio.us social bookmarking website: http://del.icio.us/)
26. (Flickr photo sharing and tagging website :http://www.flickr.com/)
27. (Etsy handmade goods website:http://www.etsy.com/)

28. Heath, C., Luff, P.: Disembodied conduct: Communication through video in a multi-media office environment. In: CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press (1991) 99–103
29. Dourish, P., Bellotti, V.: Awareness and coordination in shared workspaces. In: CSCW '92: Proceedings of the 1992 ACM conference on Computer-supported cooperative work, New York, NY, USA, ACM Press (1992) 107–114
30. Dourish, P., Bly, S.: Portholes: supporting awareness in a distributed work group. In: CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press (1992) 541–547
31. Gutwin, C., Greenberg, S.: Effects of awareness support on groupware usability. In: CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co. (1998) 511–518
32. Milewski, A.E., Smith, T.M.: Providing presence cues to telephone users. In: CSCW '00: Proceedings of the 2000 ACM conference on Computer supported cooperative work, New York, NY, USA, ACM Press (2000) 89–96
33. Tang, J.C., Yankelovich, N., Begole, J., Kleek, M.V., Li, F., Bhalodia, J.: Connexus to awarenex: extending awareness to mobile users. In: CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press (2001) 221–228
34. Nardi, B.A., Whittaker, S., Bradner, E.: Interaction and outeraction: instant messaging in action. In: CSCW '00: Proceedings of the 2000 ACM conference on Computer supported cooperative work, New York, NY, USA, ACM Press (2000) 79–88
35. Nardi, B.A., Whittaker, S., Schwarz, H.: Networkers and their activity in intensional networks. Computer Supported Cooperative Work **11**(1-2) (2002) 205–242
36. Nardi, B.A.: Beyond bandwidth: Dimensions of connection in interpersonal communication. Computer Supported Cooperative Work **14**(2) (2005) 91–130
37. Steve Alexander: Sales of Laptops Zip Ahead of Desktops (unknown) http://www.wirelessnewsfactor.com/perl/story/21861.html.
38. Handel, M., Herbsleb, J.D.: What is chat doing in the workplace? In: CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work, New York, NY, USA, ACM Press (2002) 1–10
39. Gaim website (2006)
40. Spreitzer, M., Theimer, M.: Providing location information in a ubiquitous computing environment (panel session). In: SOSP '93: Proceedings of the fourteenth ACM symposium on Operating systems principles, New York, NY, USA, ACM Press (1993) 270–283
41. Hudson, S.E., Smith, I.: Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In: CSCW '96: Proceedings of the 1996 ACM conference on Computer supported cooperative work, New York, NY, USA, ACM Press (1996) 248–257
42. Iachello, G., Smith, I.E., Consolvo, S., Abowd, G.D., Hughes, J., Howard, J., Potter, F., Scott, J., Sohn, T., Hightower, J., LaMarca, A.: Control, deception, and communication: Evaluating the deployment of a location-enhanced messaging service. [52] 213–231
43. Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J., Powledge, P.: Location disclosure to social relations: why, when, & what people want to share. In: CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM Press (2005) 81–90

44. Lederer, S., Mankoff, J., Dey, A.K.: Who wants to know what when? privacy preference determinants in ubiquitous computing. In: CHI '03: CHI '03 extended abstracts on Human factors in computing systems, New York, NY, USA, ACM Press (2003) 724–725
45. Patil, S., Kobsa, A.: The challenges in preserving privacy in awareness systems. Technical report, Institute for Sofware Research, University of California at Irvine (2003) http://www.isr.uci.edu/tech_reports/UCI-ISR-03-3.pdf.
46. Wisneski, C., Ishii, H., Dahley, A., Gorbet, M.G., Brave, S., Ullmer, B., Yarin, P.: Ambient displays: Turning architectural space into an interface between people and digital information. In Streitz, N.A., Konomi, S., Burkhardt, H.J., eds.: CoBuild. Volume 1370 of Lecture Notes in Computer Science., Springer (1998) 22–32
47. Letchner, J., Fox, D., LaMarca, A.: Large-scale localization from wireless signal strength. In Veloso, M.M., Kambhampati, S., eds.: Proceedings of The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, AAAI Press AAAI Press / The MIT Press (2005) 15–20
48. Shaffer, J., Siewiorek, D.P., Smailagic, A.: Analysis of movement and mobility of wireless network users. [51] 60–69
49. (Microsoft live local:http://local.live.com/)
50. Gellersen, H.W., Want, R., Schmidt, A., eds.: Pervasive Computing, Third International Conference, PERVASIVE 2005. In Gellersen, H.W., Want, R., Schmidt, A., eds.: Pervasive. Volume 3468 of Lecture Notes in Computer Science., Springer (2005)
51. Mase, K., Rhodes, B., eds.: Ninth IEEE International Symposium on Wearable Computers. In Mase, K., Rhodes, B., eds.: ISWC, IEEE Computer Society (2005)
52. Beigl, M., Intille, S.S., Rekimoto, J., Tokuda, H., eds.: UbiComp 2005: Ubiquitous Computing, 7th International Conference, UbiComp 2005. In Beigl, M., Intille, S.S., Rekimoto, J., Tokuda, H., eds.: Ubicomp. Volume 3660 of Lecture Notes in Computer Science., Springer (2005)