An Ecosystem For Learning and Using Sensor-Driven IM Status Messages

Donald J. Patterson, Xianghua Ding, Samuel J. Kaufman, Kah Liu, Andrew Zaldivar
University Of California at Irvine, USA
{djpatter, dingx, kaufmans, kliu, azaldiva}@uci.edu

Author Keywords
context-aware computing, instant messenger, awareness

ACM Classification Keywords
H.5.m INFORMATION INTERFACES AND PRESENTATION (e.g., HCI) (Miscellaneous)

ABSTRACT
In this article, we present a prototype system that automatically infers users’ place, activity and interruptibility from sensors on their computer and then reports this information to their instant messaging buddies. Rather than trying to interpret users’ context independently from their needs however, we approach the problem as one of supporting the user in repeating their own labeling behavior in similar situations. We present additional applications that support our view that sustained status setting behavior depends on a complete communications ecosystem that provides for easy status entry, a variety of perception channels, and intrinsic motivation.

INTRODUCTION
The Internet is radically transforming itself in response to new perspectives that highlight ideas of community and group content creation and push technological dominance to the background. This approach, sometimes called Web 2.0, drives many high-traffic websites such as video-sharing site YouTube, news-voting site Digg, and knowledge-capture site Wikipedia. These sites have modest technological innovations, like anonymous content editing, that are easily duplicated. A lack of technological dominance doesn’t threaten their success, though. They are leaders in their market because they are the sites that people actually use. The community who is creating and consuming the content is their most valuable asset.

These sites generate content through active user engagement. A person uploads a video to YouTube, submits and/or votes for news stories at Digg, and edits articles on Wikipedia. People engage in these active behaviors in order to enhance, support and maintain a community of similarly engaged people. In contrast, it is much rarer to find sites that make use of the content that people generate passively. Implicit voting is the notable exception. By watching a video, people passively cast a vote for the most popular video on YouTube.

In our system, Nomatic (“Nomad” + “Automatic”) [12], we are exploring one way in which this passive style of content creation can be amplified and leveraged. Our focus is on status messages. These short bits of text are usually created by users in the context of small communities of people who are monitoring each other for playful or work related distributed coordination. Status messages appear in instant messaging (IM) clients as short customizable phrases like “at lunch” or “out of the office.” Commercial services also provide facilities for communicating status without IM (e.g., Facebook, Twitter and Jaiku).

Structure and Sensors
We incorporate two unique ideas that make our approach to status messages useful for passive content generation. First we focus on mobile, structured, user-entered status. We are particularly interested in status that people use to describe their activities when they are engaged in specific tasks while out and about in the world on laptops and phones, but not on desktop computers. Our user-interfaces encourage status to be described through fields labelled “place”, “activity”, and “other”. We find this approach compelling because it aligns with existing practices of mobile users [11], reflects more directly on aspects of the physical world, and enables algorithms to make more assumptions about status content than would otherwise be appropriate.

Secondly, we rely on sensors that are built into commercial-off-the-shelf mobile platforms to help contextualize these status messages. By pairing structured status messages with sensor information such as wifi access points, ambient light levels, and accelerometer readings, a user’s status choice can be predicted. We use machine learning algorithms to recognize situations in which sensor readings are similar to past situations and give the user fast access to entering those status messages again.

An Ecosystem
A viable passive content generation system requires an ecosystem that supports the user’s primary interest in using status messages. In order to design such a system we build on Fitzgerald’s insights [3] which identified provision and perception as key aspects of an effective status ecosystem. To these two ideas we add motivation as a third factor that is necessary for status to be used as passively generated content. Provision draws attention to the fact that users must be able to easily and effectively describe their current status. Perception recognizes that other people must be able to view status for awareness to occur. By adding the idea of motivation, we emphasize that for a user to keep status accurate and for a computer to be able to interpret the status as content, it must be embedded in a needed and well understood task.
Our theoretical orientation comes from an idea described by Dourish as embodied interaction [2]. This approach emphasizes the fact that pervasive computing is situated in a social and physical world in which context is dynamic and constantly renegotiated between people. We view status messages as digital probes into this negotiation. Rather than giving people a tool to “geo-tag” the world, and then making them guess about how such tags would ever be used, we approach the problem as one of observing users as they conduct their everyday activities in the physical world and then using that information as a building block for improving computational services.

PROVISION

Providing easy ways for a user to enter their current status is important for making the status ecosystem viable. Most current IM systems provide options at two extremes. At one extreme users can enter completely custom status messages. Custom status provides the most nuanced control over the presentation of a user’s context, but keeping the messages up-to-date requires the user to return focus to the IM client repeatedly. At the other end of the spectrum are status indicators that are simply reports of raw sensor information whose interpretation is left open to the audience. Typically IM clients provide this service as an “idle” indicator that is based on a lack of keystrokes or the invocation of a screensaver. The first extreme produces very appropriate contextual information. The second extreme communicates accurate sensor cues which require interpretation. Our approach is to create a hybrid of these two extremes that, like sensors, support rapid easy updates but like custom messages, also provide an interpretation of that information (see Figure 1).

Related Work

Our machine-assisted approach uses sensors to provide rich descriptions of context through mobile status messages. In the process this problem begins to look a lot like related research in place, activity and interruptibility detection. By combining all three tasks in one application, however, we are able to provide a rich set of data for human observers to interpret. For example, knowing that someone is in a classroom provides a great deal of information about what activities a person may be engaged in. Knowing that someone is in a classroom and teaching provides a lot of information about how interruptible they are as well. To the best of our knowledge no other researchers are investigating how these various aspects of context can be linked to help inform the others.

Hightower et.al., framed a portion of our problem as the “Position” to “Place” problem [7]. In that work, the problem was translating the exact unambiguous sensor data from GPS streams (or wifi-based localization) into a description that was more reflective of what the place was called. Their research brought up a number of relevant challenges with such a mapping. Principle among them are the fact that there is a many-to-many mapping of positions to places and it is difficult to argue for what the correct place name is. We acknowledge these concerns by changing the focus of the problem from picking the correct place name based on sensor data, and instead attempt to recommend an appropriate place name, as validated by the user’s selection of our recommendation.

This work is also closely related to activity recognition, in which many researchers are attempting to develop techniques to label a user’s current activity from sensor streams. Much of this work attempts to characterize a user’s sensor stream as being generated from 1 of N exclusive activities to support underlying reasoning algorithms (e.g., [13]). We are unable to precategorize which activities our users will be engaged in beforehand, so we take a different approach to activity recognition in which we simply attempt to suggest the same activity to a user when they are in the same situation (as measured by the available sensors). This permits the user to have complete control over how they wish to describe their activity, but limits the types of machine learning approaches that are available to us.

This work is also closely related to work in interruptibility detection from sensors (e.g., [1, 8]). However, for our desired application domain we need a more nuanced approach to interruptibility that gives users more flexibility than a numeric scale between interruptible and not-interruptible. On the one hand, an on-call doctor in a restaurant may be interruptible for an emergency, but not interruptible for a billing question. When the domain of interest moves away from “the office” the question of interruptibility is much harder to define. Our approach instead is to push that decision into the

![Figure 1. Nomatic uses sensors to provide IM buddies with user-generated contextual cues. These semantic interpretations are easier to keep up-to-date than custom status lines and are easier to interpret than raw sensor data.](image-url)
social sphere. We attempt to learn a semantic description of the current sensor stream so that people can make a decision about whether or not it is appropriate to initiate a communication: If a caller knew that their callee was in a restaurant, they could self-censor themselves appropriately.

As one would expect, we are also not the only researchers that are working with context-enhanced IM systems. Other researchers are looking at how people respond to context-enhanced IM in regard to privacy (IMBuddy [9]). We are less concerned with privacy because we don’t automatically broadcast sensor readings, we only broadcast status information that is predicted from sensor readings after the user has accepted them. Furthermore our status information mimics the same language that the user used in a similar situation in the past, so users maintain control over their digital presentation. For example, if a user is in a coffeeshop working and reports being “at work”, we repeat that behavior even though an external location ontology would be unlikely to label Starbucks as “work”.

The Awarenex system uses sensor modules to help support fluid conversation openings and closings [14] which is similar to Fogarty’s work on interruption management [4]. We differentiate ourselves from these researchers by focussing on accurately predicting user-generated labels from the sensors that are built into current hardware. These labels may serve similar functions in conversations, but their function is less of a focus in our research.

Prototype Design
An effective user-friendly method of entering status requires that entering the correct status be fast and easy. We prototyped such an interface using a combination of machine learning techniques and user interface design. The program flow works as follows:

- **Steady State:** While the user is working, our stand-alone program monitors the sensors on the computer and displays the current status in a small window.
- **Change Status:** When the user presses a button to change their status the interface reads the sensors on the computer, and uses a machine learning algorithm to make predictions for an appropriate status.

  - **Validate Predictions:** The predictions are shown in a selection box (see figure 2, left). Based on earlier results [11] that indicated that 71% of all custom status messages in mobile IM are used to describe, place, activity or availability, we offer the user a template to enter place, activity or a free-form label called “other”. In the best case the algorithm is correct and users only have to accept the guesses by hitting “Change”. Otherwise they can pick from the ranked list of guesses or, ultimately, type in a new status setting.

  - **Accept Prediction:** When the user accepts the new status, the interface again sweeps the sensors on the computer, pairs the sensor readings with the current status and updates a local database that is used by the machine learning algorithm for future training. See the dataflow in figure 1.

Evaluation Methodology
In evaluating the machine learning aspect of this interface we asked 14 mobile laptop IM users to use the status interface for 3 months and followed up during and after the study with short interviews regarding their impressions after usage. We collected 7154 status lines paired with sensor data that yielded an average of 5.7 status changes a day per person. From this information we evaluated several classification strategies as shown in table ??.

For each of these machine learning methods we tested 7 different prediction tasks. The features consisted of available sensors, \( \hat{s} = \{ \text{day of the week, time of day, local and remote IP address, the wifi AP MAC address and SSID, the currently active process, the number of displays connected} \} \).
to the computer and their resolutions, whether the mobile device was plugged in or not, 3-D accelerometer readings, ambient light, the volume setting on the computer, and the number of mouse clicks per second). Depending on the task, the following were used as features or classification targets, place status, p, activity status, a and the other status, o. From these features we calculated \( P(p|\hat{s}), P(a|\hat{s}), P(a|p), P(a|p, \hat{s}), P(o|\hat{s}), P(o|p, a), P(o|p, a, \hat{s}) \)

### Analysis
By analyzing the results on our dataset we see that users have strong patterns of repeated status setting behavior in mobile IM (see table 1). Based on the results from the single most likely classifier for place, \( P(p|\hat{s}) \), we can see that 49% of the time users report being in their most frequent location, 32% of the time they are doing their most frequent activity and use a single unique other status 71% of the time. Although these numbers are high from a temporal perspective they mask the variety of ways that people described their places. The average number of unique places reported by a person was 16. The most common places were custom variations on “At home” and “At work”, but included things as varied as “at the beach” and “in my car”. The most common choice for activity status was to leave it blank, but that was closely followed by variations on activities such as “writing”, “sleeping”, and “working” and included “doing dishes” and “eating a chocolate muffin”. Other status was most often blank and after that was hard to categorize, with examples such as “George Lopez!!”, “Don’t bother me”, “sick” and “Life’s good” being typical.

Across the board, support vector machines (SVMs) offered the best overall classification accuracy. SVMs were able to predict the same place name that the user picked in 94% of the cases based on the sensors available, \( P(p|\hat{s}) \). Similar results were found when predicting a user’s activity, \( P(a|\hat{s}) \). The large gains made by the Decision Stump algorithm over the Most-Likely algorithm in predicting place, \( P(p|\hat{s}) \), confirms the intuition that a few sensors contain a lot of information about place (e.g., wifi AP), but that to do the best job of picking a nuanced place name requires algorithms that leverage a wide variety of non-location specific sensors. If you only knew how a person described their place, SVMs had a 45% chance of correctly guessing the associated activity, \( P(a|p) \), which is a 13% improvement over just picking the most likely activity as a default. But adding the user’s choice of place names on top of the sensors available \( (P(a|\hat{s}) \rightarrow P(a|p, \hat{s})) \) had negligible effect on improving activity prediction (93% → 94%). Finally in predicting the other status field, having knowledge of place and activity allowed SVMs to achieve 81% accuracy, \( P(o|p, a) \), but the sensor information by itself was sufficient to improve that to 96% accuracy. As was the case with predicting activity, the additional interpretation by the user of the current place and activity \( P(o|\hat{s}) \rightarrow P(o|p, a, \hat{s}) \) was of negligible benefit (96% → 97%).

Because keeping status accurate is also a matter of keeping the status up-to-date, we investigated strategies for interrupting the user at moments when it appeared that status might be changing. Our goal was to identify the most helpful times to interrupt a user. We programmed Nomatic to observe the sensor stream and look for a variety of different changes and when they occur, to pop up the status change interface. After asking the user to set their status, we also asked them to rate how helpful the reminder was. The interface, triggers and results of the strategies are shown in figure 2.

From these data we can see that users generally find it helpful to be prompted to change their status simply based on elapsed time, whether that occurs after startup or after the user has been using the computer for a while. One effect that generated this result is that when a user takes a laptop out of standby, our interface is triggered. Intuitively this seems like a good time to get input from the user and our data confirms this. In contrast, interrupting the user during network change events is not as clearly helpful. Although this seems to be a good indicator of a change in context related to mobility, our users experienced many situations in which their wifi connection rapidly switched between several competing access points while they were not changing status. Similarly intermittent internet network problems made user’s local and remote IP address change without a corresponding change in the user’s location resulting in high annoyance scores for those triggers.

### PERCEPTION
Using a status line to communicate aspects of your context is not helpful unless the ecosystem also provides for perception of status by others. We enabled Nomatic to interface with several 3rd-party status broadcast services. Users can optionally choose to have their status reported to local IM clients (e.g., pidgin, iChat, Adium) or popular microblogging services (e.g., Twitter, Facebook). Each of these ser-

| Algorithm       | \( P(p|\hat{s}) \) | \( P(a|\hat{s}) \) | \( P(a|p) \) | \( P(a|p, \hat{s}) \) | \( P(o|\hat{s}) \) | \( P(o|p, a) \) | \( P(o|p, a, \hat{s}) \) |
|-----------------|-------------------|-----------------|-------------|-----------------|-----------------|-----------------|-----------------|
| Most-Likely     | 49%               | 32%             | 32%         | 32%             | 71%             | 71%             | 71%             |
| Decision Stump  | 64%               | 38%             | 35%         | 38%             | 74%             | 73%             | 73%             |
| K-NN            | 75%               | 50%             | 43%         | 53%             | 74%             | 77%             | 77%             |
| Naive Bayes     | 85%               | 72%             | 44%         | 73%             | 78%             | 79%             | 82%             |
| Decision Tree   | 85%               | 63%             | 46%         | 64%             | 81%             | 81%             | 84%             |
| Boosted Stumps  | 91%               | 76%             | 44%         | 80%             | 85%             | 80%             | 91%             |
| SVM             | 94%               | 93%             | 45%         | 94%             | 96%             | 81%             | 97%             |

Table 1. The probability of various machine learning algorithms correctly predicting place(p), activity(a), and other (o) status based on combinations of status and sensor (s) readings. Best performing combinations that are statistically indistinguishable \( (p < 0.05) \) are shown in bold.
Figure 3. Nomatic*Bubbles displaying 30 days worth of data. The largest brown ring at the center is the UCI campus-wide SSID. The blue rings are wifi access points labeled with place names using our status interface. The size of the blue ring indicates the number of people who have visited the location, while the border thickness indicates how often it is visited (in days). The brown ring clusters wifi access points that have the same SSID. Its position is determined by its size (the number of wifi hotspots associated) and the recency of the last visit. The colored dots indicate the most recent places from which users have reported.

Vices have mechanisms for broadcasting status to a social network of various types so that status can be used to negotiate appropriate responses to context.

Thinking about perception motivated us to make a particular design decision in our tool that we would not automatically change status for a user. Because status communicates potentially sensitive information, we feel it is important that a user trusts our tool not to broadcast an inappropriate status. We support that by requiring the user to stay in the loop and to press at least two buttons to initiate and then accept a change their status. This exposes the relationship between accuracy and user attention that Nomatic is trying to minimize, but we hypothesize that remote viewers will find context status more useful if they can be confident that it was confirmed before it was set. If buddies begin to doubt that status is accurate they are likely to ignore it, and render it useless.

Nomatic*Bubbles

Another way of providing perception of status is through communal awareness displays. We designed such a display called Nomatic*Bubbles which visualizes the information that members of a small community can provide through our status interface. Because our software uses hardware that comes with existing computing platforms, rather than custom infrastructures, we are forced to address new ways of representing location in semantically meaningful ways when our users move outside of research environments. Instead of using geographic maps, Nomatic*Bubbles depicts historical and aggregate traces of participants’ whereabouts in an abstract and ambiguous manner, as a way to convey the more useful context information that status provides, while attempting to stay away from connotations of “tracking” or “monitoring” that raise privacy concerns.

Figure 3 shows a snapshot of Nomatic*Bubbles with thirty days’ worth of data. In this visualization, the layout is dynamically determined by users’ collective interactions with the wifi infrastructure, it depicts historical traces of people’s whereabouts and instead of using explicit icons for people, it uses different colors to distinguish real-time context of individuals. Therefore, participants and the people who are engaging with them and the display, gradually learn how to interpret the information. People who are casual observers have limited insight into the details of a particular person’s context.

Preliminary feedback of this display, which has been operating in the elevator lobby of our building for more than a month, has been mixed. Users and viewers of the system have revealed both an intuitive understanding of the abstract representations on one hand and confusion on the other. Generally users’ ability to interpret the visualization is correlated with their engagement with the community being visualized, which is a success of our design goals. While viewers cannot interpret all traces, they can gain general impressions of the overall activities of the community, and can easily recognize and interpret some of the participant’s data. Secondly, the physical setting of the deployment limits its effectiveness. Somewhat counterintuitively, the elevator lobby is too transitional a space and wait times are too short for people to be able to develop an familiarity with what is going on in the visualization.

MOTIVATION
Maintaining and monitoring status requires effort on the part of a user. The ability to simply log status motivates some users to keep status accurate. Such record keeping is useful for managing billing records and for encouraging self-reflection. Additional motivating applications help to further increase the benefits without additional effort and subsequently improve the ecosystem for passive content creation.

**IM Interruption and Embarrassment**

Simply adding to ability for others to see your status creates awareness, a combination of provision and perception. This has been thoroughly documented as having intrinsic value to support and improve distributed group work (e.g., [6, 15]). Studies have indicated that 13% of all pre-mobile IM dialog was simply related to negotiating availability [5]. As individuals are increasingly always “online”, IM has moved onto mobile platforms and interruptions appear to be getting worse. More recent studies of mobile IM users have reported that 43% of users are actively using strategies to manage interruptions, and that as many as 7% of mobile IM users have stopped using IM at one point or another due to the distractions that it causes. 80% of undergraduate mobile IM users report receiving embarrassing IM’s because of the semi-public visibility of their laptop screens [11]. So if, as we hypothesize, machine-assisted status messages do effectively mitigate interruptions and alert users about the social context of a buddy before they initiate an unintentionally embarrassing, then that is a significant motivation for keeping status up to date.

**CATDL: a Context-Aware To-Do List**

Another way that status can be used is to incorporate it with a context-aware to do list. By reordering a long list of tasks to focus attention on those that are relevant for the current situation, using a process that the user is already engaged in, we make the maintenance of context even more valuable. This approach enables a grocery list to appear on a mobile device upon entering a grocery store and a list of work tasks to naturally shift to a list of home tasks at the end of a work day.

We have implemented such a system as a locally hosted web application called Nomatic*CATDL which combines the status from Nomatic with an online to-do list manager called Remember the Milk1. When a user wants to see their current to-do list, they open the web application which establishes a local connection to the Nomatic status setting tool and a remote connection to Remember The Milk. The task list is contextually sorted in the browser and presented as a lightweight user interface suitable for a mobile phone.

**Geographic Overlays**

Although the display of Nomatic*Bubbles, mentioned earlier, was highly abstract there are also times when it is appropriate to more explicitly map the geographic location of context. In public or professional spaces people often explicitly reveal or are required to reveal the task in which they are currently engaged in. In such situations it may be effective to explicitly represent status as a function of position. We are exploring this idea through the design and study of a touch-screen directory kiosk for a large academic building. Existing kiosks help people find location according to facility managers descriptions (such as building names, room numbers, etc), but don’t reveal colloquial names for places, such as a familiar name for a class taught in that building (e.g., “Physics for poets”).

We have prototyped a display that overlays a traditional building map with historical and real-time status information (see Figure 4). In this design a user can select a temporal range of data that they would like to visualize at the top. This selection populates three drop-down lists on the lower left with data that other users have previously entered through the Nomatic tool. This building has been calibrated to support localization by wifi fingerprinting, so status can be associated with a particular geographic location. To support anonymity, only those status entries that are duplicated by more than one person are displayed. The user of the display can then select a place, activity or other status from the drop down list and locations at which that status label has been used are highlighted with colored circles on the map. We are currently in the process of studying how the enhanced design of the kiosk affects user’s interaction with the directory display.

**CONCLUSIONS AND FUTURE WORK**

As we continue to explore the space of possibilities and identify how to innovate the existing technologies to best support user’s needs we see many opportunities. There is interesting work to be done in developing a more sophisticated understanding of when a user would like to be interrupted to change their status. There are opportunities for developing

---

1www.rememberthemilk.com
mass-collaboration approaches for sharing sensor and status information across users. There is the possibility of identifying which sensors are the most valuable for setting status as a way of advising future hardware construction. This also includes creating abstract virtual sensors based on a user’s calendar, the statuses of buddies, or any number of digital probes that may influence the mood of a user.

By simply attaching sensor data to the status information that users are entering in IM we can create a rich ecosystem of context-aware applications that benefit the end user. At the most basic level keeping status up-to-date helps to mitigate the increasing problem of interruptions in mobile communications, but also has the potential to enable many other uses of the data. This ecosystem must be kept in balance by supporting the user’s ability to provide the status information, supporting other users ability to see the status information, and providing motivating reasons for why either of them should want to make the effort to keep it accurate.

REFERENCES


