Exploiting Context Data Fidelity for Enhanced Privacy and Energy Savings

Angela B. Dalton, Carla S. Ellis, and Abhijit Vijay
Duke University
Department of Computer Science
`angela,carla,abhijit@cs.duke.edu`

May 7, 2004

Abstract

Mobile computing devices with access to wireless networking and inexpensive sensors offer the promise of many exciting new context-aware applications. Unfortunately, the promise of ubiquitous context-aware applications has not been realized as rapidly as hoped. Two major barriers to the widespread adoption of this technology have been (1) limited battery lifetime of the mobile devices carried by users or the remotely deployed sensor nodes instrumenting the environment and (2) perceived threats to privacy by what might be interpreted as surveillance. Based on our experience with context-aware systems, we propose a model of data fidelity that represents the tradeoffs related to energy consumption and privacy in context-aware applications. The model identifies the dimensions of data fidelity as the capture, the persistence, and the dissemination of sensor and context data.

1 Introduction

Ubiquitous computing systems comprised of wireless networks and inexpensive sensors offer the promise of a broad range of valuable context-aware applications. These applications would deliver high value services that are specifically tailored to the needs and desires of individual users based on their physical and personal context. Much of the context information needed by such services can be captured by augmenting the mobile devices routinely carried by users and instrumenting the environment with various sensors.

A commonly cited example of a ubiquitous system allows a user to navigate through an unfamiliar city with the aid of a location-aware service. Another attractive application is a cellular phone that can automatically sense situations in which forwarding to voicemail is more appropriate than ringing, thus releasing the user from the burden of manually managing his phone's modes.

Unfortunately, the promise of sensor-based context-aware applications has not been realized as quickly as hoped. Two major barriers to the widespread adoption of this technology have been

1. limited battery lifetime of the mobile devices carried by users or the remotely deployed sensor nodes instrumenting the environment.

2. perceived threats to privacy by what might be interpreted as surveillance.

Eliminating these two major issues of ubiquitous computing is key to accelerating the acceptance of sensor-based context-aware systems. To be effective and widely adopted, context-aware systems must not present users with yet another device to manage or cause them to constantly worry about recharging batteries and about the risks involved in sharing information.

In this paper, we propose a conceptual model of data fidelity and explore how to effectively use fidelity to achieve the requirements for energy efficiency and privacy assurance in context-aware,

ubiquitous systems. We use the term "fidelity" to reflect the precision and faithfulness of a data representation relative to some reference object. We first describe the context-aware systems we use as the basis for our research. Then we explain our fidelity model and show how use of the fidelity model benefits our own context-aware systems in reducing energy consumption and providing increased privacy assurance.

2 Our Experiences

The FaceOff Project uses low-power sensors as tools in the service of OS-based energy management for mobile computers. In FaceOff, we consider sensors providing information from which to infer user intention and user context as it affects energy management of the display, capturing the direct dependency that looking at the screen suggests a need for it to be illuminated. Intuitively, this is attractive as a more direct indication of the user's need for display power consumption than the keyboard and mouse input events used in traditional timeout-based strategies.

The FaceOff design [3] consists of three main components: image capture, face detection, and display power state control. FaceOff periodically wakes up and calls the image capture component. The image capture mechanism obtains a still image from a camera and sends the image to the face detector for analysis.

Face detection can be a computationally intensive technique. However, it can be specifically optimized for the simplified problem of detecting an upright, frontal face of an approximate size indicating the presence of a user looking at the display. Currently, the face detection module consists of a skin color detector that looks for a large central area of skin color in the image. Skin color detection was selected as a fast and fairly simple method for the initial prototype. Thus, the information used is low fidelity yet it proves surprisingly effective for our purposes[1]. The face detec-

---

[1]More accurate fast face detection methods exist and are part of our longer term plans for the FaceOff project that call for evaluating whether the higher fidelity information might be useful.

tor returns the Boolean value of true if a face is detected and false if no face is detected. The display power is controlled using ACPI (www.acpi.info) to change the video device power state, setting it to a sleep mode when no face is detected.

We built the initial FaceOff prototype on an IBM T21 Thinkpad running Red Hat Linux. The camera is a color Logitech QuickCam 3000 web cam that connects via USB to the laptop with an average measured power consumption of 1.5W. Measurements of energy consumption using the prototype system indicate the promise of significant energy savings from this type of context-based display power management scheme. It

Figure 1: Low fidelity image used by FaceOff – an image with skin colored pixels identified in black.

is not unusual today for mobile devices to be equipped with integrated low resolution cameras. Privacy concerns arise for many users when images are captured without their explicit knowledge and consent. In the case of the FaceOff system, we address these concerns by not storing images and performing minimal processing to determine only the presence or absence of skin blobs. Figure 1 shows a skin blob detected in an image with FaceOff. Both of these techniques also contribute to a lower energy overhead for the FaceOff system, improving its utility as a display management system.

We added an x10 wireless motion sensor to the prototype for experimentation with alternative or additional sensors and context information. By

using the motion sensor to provide a lower fidelity source of context data we are able to save more system energy overall by eliminating even the overhead of the camera and face detection computation during long periods of time with no motion present.

Many mobile devices today are equipped with a variety of low power sensors. The cameras, light sensors, microphones and wireless interfaces incorporated into many devices can be used to collect context information about the user's behavior. We are exploring other ways in which to use context information from such sensors to reduce the device energy consumption.

Location-aware systems are an important subclass of context-aware systems. Our group recently developed a client-side location tracking system using signal strength readings from multiple 802.11 access points [14]. This approach leverages existing infrastructure in the building. Our system, Uhuru, provides a proof of concept of the ability to determine location information in three dimensions, across multiple floors of a building. In normal operation, Uhuru does a simple table lookup against the database of signal strength values and physical locations that it has built during a previous initialization phase and reports the physical location corresponding to the best match. The lookup is done by computing the Euclidean distances in signal space. Fidelity issues arise in that signal strength data is not precise or consistent.

In order to improve the accuracy of the location tracking system, we devised and implemented a limited history-based algorithm. This algorithm is based upon the premise that the mobile user cannot switch from one set of coordinates in the physical space to another totally arbitrary location, from one instant of time to the next. Clearly, only recent history is relevant to this method and the longer the period between subsequent readings, the less faithful to the underlying assumption of continuous user movement the history becomes. Experiments showed an improvement in accuracy with this limited history. Both the limited history retention as well as the client-side processing of signal strength data in Uhuru may be

considered by users as features that enhance privacy.

Our experience with the FaceOff and Uhuru systems suggest some common characteristics when viewed in terms of how the context data are handled. This motivates our approach to jointly address the energy and privacy problems: to focus on sensor data fidelity and related operations that affect it. We believe that many context-aware applications should be thought of in terms of the role data fidelity plays in their overall utility, both due to its effects on privacy and energy consumption.

3   The Fidelity Space

The adaptation of data fidelity has been one of the most important techniques employed in energy-aware software design. Reducing fidelity can reduce the amount of work required and, consequently, the energy consumed to deliver a service. Reduced fidelity of sensor and context data may also be seen to disclose less personal information.

The need to understand how to adjust data fidelity to obtain appropriate levels of privacy and energy consumption motivates the definition of a "fidelity design space." This yields a model to understand design tradeoffs as they relate to energy consumption and privacy.

We are formulating a fidelity space that is comprised of dimensions relating to the *capture* of context data, the *storage* of the data, and its *dissemination*. Our first dimension concentrates on the initial data capture and immediate processing of the sensor readings or from the context source. Characteristics of the data being captured are important since they set an upper bound on the information content that subsequent processing steps may inherit. For example, in FaceOff, either a high resolution image taken by a camera or a narrowly focused infrared sensor may be used to indicate the presence of a person sitting in front of a display to an accuracy that satisfies the requirements of our particular application. While the fine-grain image data may be minimally processed to serve the immediate purpose (e.g., detection of skin color only), the user may not feel

confident that the image, once it has already been collected, will not be used for identification, as well. By contrast, the infrared data may not be perceived as a similar threat to privacy. This example highlights one of the challenges of characterizing the collected data – we must consider different *types* of data as competing alternatives in achieving a specified functionality. While it may be straightforward to compare a low resolution image and a high resolution image, it is more difficult to consider image data and infrared data on a single scale. Precision of the data is type-specific (e.g., resolution for images, sensitivity for pressure sensors, etc). Capture also represents a lower bound on the energy usage of acquiring the raw data. Obfuscation is a technique for addressing the privacy issue, but it consumes additional energy cost in post-processing the original finer-grain data.

The second dimension concerns the *storage* of the collected context data. Storage obviously has an impact on energy consumption. Storage is also the mechanism that determines the persistence properties of the data and potentially associates it with a temporal aspect. Logging the history of sensor readings has different implications on what can be done with the data than having only the last instance available. There is also a granularity parameter involving the frequency of context snapshots or freshness of the latest reading. Data points may be associated with precise timestamps or they may just be ordered. Do older data decay with time? How does persistence relate to the accuracy of the application? The mechanism provided in our location-tracking system, Uhuru, to disambiguate candidate positions illustrates how recent history can help accuracy but its usefulness degrades over time for that purpose. The potential for long-term storage raises privacy concerns in that the data can be subject to unanticipated uses after the initial cost/benefit bargain is assessed by the user.

The third dimension involves the *dissemination* of the data. This physically relates to the networking aspects of the application and affects the energy consumed in communication. Dissemination implies potential loss of control over the data.

Transferring the data over the network may not be intended as explicit data sharing, but it may affect control over the site of data storage (e.g., client-side vs. server side). For example, location information saved as a track on the user's GPS-enabled handheld device may be considered private, unless it is voluntarily disclosed. Voluntary communication is obviously essential in many context-aware applications. An example is when the desired functionality is to relate one's location to that of another person. On the other hand, when location information is recorded by the infrastructure (e.g., installed readers of passive RFID tags carried on the user), the inherent lack of control may be viewed as a privacy threat. Issues such as data aggregation within a sensor network can be viewed in a fidelity framework.
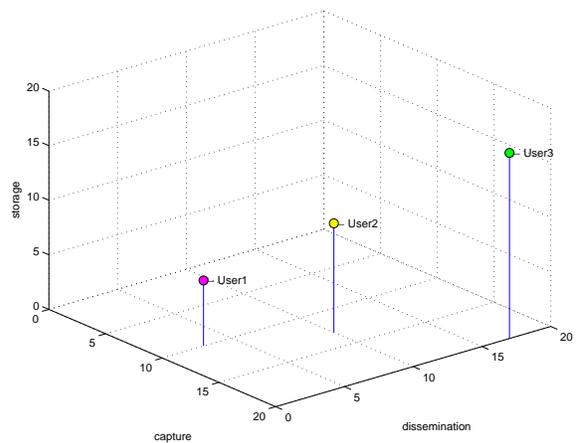

Figure 2: Preferences of 3 users located within the fidelity space.

Figure 2 illustrates how different users' preferences might be located within our fidelity space. The location within the space corresponds to the value between 0 and 20 for each axis, where 0 represents the lowest fidelity and 20 represents the highest. User 1 is highly concerned with protecting her privacy and wants to be sure that her context information is not stored or disseminated. User 3 is very diligent in charging her mobile device batteries and is not worried much about energy constraints. She is more interested in deriving maximum benefits from disseminating her context data to what she views as userful services

and maintaining a complete history of her data. User 2, while not extremely concerned with privacy of context information, does not want it too widely disseminated because she sees some risk in losing control over the information. Also, she is not as likely to keep all of her batteries charged and is concerned about the lifetime of her mobile devices. For the same reasons she selects a middle range for the storage of her context information.

Combinations of sensor data may present both risks or benefits that individual sensor values do not. In FaceOff, the camera and the motion sensor each suggest the presence or absence of a user in front of the display. In this case, they play different roles and, in combination, provide better energy savings. In other cases, a combination of several very low-grade sensors may imply the result that a single high-grade sensor would capture. Understanding how sensor data correlate is an important aspect of the privacy risk analysis.

4   Related Work

Even as only an intuitive notion, fidelity adaptation has proven useful in mobile/wireless battery-powered systems. For example, Flinn and Satyanarayanan [4] demonstrate the impact on energy savings of lowering data fidelity in several applications running under the Odyssey system. The term "fidelity" is informally used to convey a variety of adaptations including various levels of lossy compression for images and videos, feature selection and cropping in a map server, smaller dictionaries used to process utterances in speech recognition, and reduced window sizes for display. Such fidelity adaptation may entail conversions of data format (e.g., from gif to jpeg, from image to text) that are challenging to compare.

Castro and Muntz [2] use a definition of the "quality of information" that includes both a measure of accuracy and a measure of uncertainty. The accuracy of conglomerated sensor data is the maximum probability that a hypothesis is true given certain evidence. This represents the predictive value of the sensor data. The goal of their system is to find the best sensor configuration in a smart room to perform a specified data collection task. They identify the sensors that give an acceptable level of accuracy while minimizing cost where cost is the total resource usage cost for the sensors in that configuration.

Techniques related to data fidelity have also been proposed to improve privacy of systems, however the methods used to anonymize data often require increased computation, which would be contrary to the goal of reduction in energy consumption. Lederer, et.al [11] suggest the use of technology to reduce the distinction between between surveillance and transaction, for example by blurring the face of specific individuals in video streams, allowing the disclosure of presence and identity to become a transaction controllable by the subject. Fidelity of a user's activity can be reduced by conveying fewer or less precise data points, as suggested in [11]. The Place Lab infrastructure [10] provides an interface to allow users control over the granularity of information revealed to external hosts for location-enhanced world wide web applications. Depending on the user's specifications, the system may reveal the user's location in terms of city, or street address, or exact room number within a building. Forgarty [5] suggests that collection of data from simple sensors can allow for very accurate prediction of some context information even when more expensive and powerful sensors are normally associated with such applications. The specific example Fogarty presents is a single microphone in an office rather than a camera being used to determine interruptibility.

Another important technique for managing privacy is using a decentralized approach to storing and analyzing data when possible. Many recently developed location tracking systems are decentralized. Place Lab [10], Cricket [13], RADAR [1], and Uhuru [14] are examples in which location determination is performed locally, giving users greater control over disclosure of their location to others. Gruteser et. al [7] address privacy in location-aware sensor networks through a distributed anonymity algorithm that is applied before service providers gain access to data. Their approach employs data processing on the sensor nodes to execute the anonymity al-

gorithm and requires substantial communication overhead making energy consumption a concern.

Hong et al [9] discuss practical guidelines to help stakeholders assess privacy risks and benefits in a context-aware system. Based on this privacy risk model, an architecture with concrete mechanisms is proposed [8] to help developers follow those practices in privacy-sensitive location services.

5 Conclusion

We have presented our insights into the role of data fidelity, specifically for context and sensor data, as it relates to the limited battery lifetime and perceived threats to privacy that hold back the widespread deployment and user acceptance of ubiquitous context-aware systems. Our experience with two context-aware systems led us to define dimensions of the fidelity design space that will assist developers and end users in understanding data fidelity tradeoffs as they relate to energy consumption and privacy. The dimensions of our model include the capture, the persistence, and the dissemination of sensor and context data. The model provides a framework for development of tools and systems support for applications exploiting fidelity adaptation for context and sensor data.

6 Acknowledgements

References

[1] Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *INFOCOM (2)*, pages 775–784, 2000.

[2] Paul Castro and Richard Muntz. Managing context data for smart spaces. In *IEEE Personal Communication*, October 2000.

[3] Angela Dalton and Carla Ellis. Sensing user intention and context for energy management. In *Workshop on Hot Topics in Operating Systems (HOTOS)*. USENIX, May 2003.

[4] Jason Flinn and M. Satyanarayanan. Energy-aware adaptation for mobile applications. In *Symposium on Operating Systems Principles (SOSP)*, pages 48–63, December 1999.

[5] James Fogarty. Sensor redundancy and certain privacy concerns. In *Workshop on Privacy in Ubicomp 2003: Ubicomp communities: Privacy as boundary negotiation*, October 2003.

[6] Deepak Ganesan, Ben Greenstein, Denis Perelyubskiy, Deborah Estrin, and John Heidemann. An evaluation of multi-resolution search and storage in resource-constrained sensor networks. In *Proceedings of the First ACM Conference on Embedded Networked Sensor Systems*, 2003.

[7] Marco Gruteser, Graham Schelle, Ashish Jain, and Dirk Grunwald. Privacy-aware location sensor networks. In *Proceedings 9th USENIX Workshop on Hot Topics in Operating Systems (HotOS)*, 2003.

[8] J. Hong and J. Landay. An architecture for privacy-sensitive ubiquitous computing. In *Proceedings of International Conference on Mobile Systems, Applications, and Services*, June 2004.

[9] J. Hong, J. Ng, S. Lederer, and J. Landay. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of Designing Interactive Systems (DIS2004)*, 2004.

[10] Jason I.Hong, Gaetano Boriello, James A. Landay, David W. McDonald, Bill N. Schilit, and J.D. Tygar. Privacy and security in the location-enhanced world wide web. In *Workshop on Privacy in Ubicomp 2003: Ubicomp communities: Privacy as boundary negotiation*, October 2003.

[11] S. Lederer, J. Mankoff, and A.K. Dey. Towards a deconstruction of the privacy space. In *Workshop on Privacy in Ubicomp 2003: Ubicomp communities: Privacy as boundary negotiation*, October 2003.

[12] Samuel Madden, Michael J. Franklin, Joseph M. Hellerstein, and Wei Hong. The design of an acquisitional query processor for sensor networks. In *Proceedings of the 2003 ACM SIGMOD international conference on on Management of data*, pages 491–502. ACM Press, 2003.

[13] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Mobile Computing and Networking*, pages 32–43, 2000.

[14] Abhijit Vijay, Carla Ellis, and Xiaobo Fan. Experiences with an inbuilding location tracking system: Uhuru. In *IEEE Int'l Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2003.