Title:
Participatory sensing

Author:
Burke, Jeffrey A
Estrin, D
Hansen, Mark
Parker, Andrew
Ramanathan, Nithya
Reddy, Sasank
Srivastava, Mani B

Publication Date:
05-05-2006

Series:
Papers

Publication Info:
Papers, Center for Embedded Network Sensing, UC Los Angeles

Permalink:
http://escholarship.org/uc/item/19h777qd

Additional Info:
Copyright is held by the author/owner(s). WSW'06 at SenSys '06, October 31, 2006, Boulder, Colorado, USA. Copyright 2006 ACM 1-59593-343-3/06/0011…$5.00. World Sensor Web Workshop, ACM Sensys 2006, Boulder, Colorado

Keywords:
participatory sensing, urban sensing, network-attested context, community-based participatory research, mobile handsets, urban planning, natural resource management, public health, cultural identity

Abstract:
This paper introduces the concept of participatory sensing, which tasks everyday mobile devices, such as cellular phones, to form interactive, participatory sensor networks that enable public and professional users to gather, analyze and share local knowledge. An initial architecture to enhance data credibility, quality, privacy and 'shareability' in such networks is described, as well as a campaign application model that encompasses participation at personal, social and urban scales. Example applications are outlined in four areas: urban planning, public health, cultural identity and creative expression, and natural resource management.

Participatory Sensing

J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava
Center for Embedded Networked Sensing (CENS), University of California, Los Angeles
3563 Boelter Hall, Los Angeles, CA 90095-1596
jburke@ucla.edu,destrin@cs.ucla.edu,cocteau@stat.ucla.edu,
{adparker,nithya}@cs.ucla.edu, {sasank,mbs}@ee.ucla.edu

ABSTRACT

This paper introduces the concept of participatory sensing, which tasks everyday mobile devices, such as cellular phones, to form interactive, participatory sensor networks that enable public and professional users to gather, analyze and share local knowledge. An initial architecture to enhance data credibility, quality, privacy and 'shareability' in such networks is described, as well as a campaign application model that encompasses participation at personal, social and urban scales. Example applications are outlined in four areas: urban planning, public health, cultural identity and creative expression, and natural resource management.

Keywords

Participatory sensing, urban sensing, network-attested context, community-based participatory research, mobile handsets, urban planning, natural resource management, public health, cultural identity.

1. INTRODUCTION

Two billion people carry mobile phones. These ubiquitous devices are increasingly capable of capturing, classifying and transmitting image, acoustic, location and other data, interactively or autonomously. Given the right architecture, they could act as sensor nodes and location-aware data collection instruments. The idea of a wireless sensor network is not new; we are already able to integrate sensing, computation and connectivity in low-power devices and embed networks of them in the physical world to collect data. Through physical proximity to what they study, multiscale operation and adaptive, autonomous observation, embedded sensor networks allow the observation of previously unobservable phenomena. From research and deployment, we know something about what distributed sensing can be used for in the sciences, industry and the military. We know much less about its function and utility in the public sphere when the components are owned and operated by everyday users. As sensors, network-connected mobile handsets will be embedded near the ultimate elusive subjects: people and their built environments. Instead of being in the hands of a central observer, these sensors are always-on and under their owners' control. Leveraging them effectively and conscientiously will require models that prioritize user participation in sensing.

Participatory sensing will task deployed mobile devices to form interactive, participatory sensor networks that enable public and professional users to gather, analyze and share local knowledge. Microphones and imagers on-board the mobile handsets can record environmental data now, while in the future other sensors will be integrated or connect wirelessly. Cell tower localization, GPS and other technologies can provide location and time-synchronization data. Wireless radios and onboard processing enable human interaction with both local data processing and remote servers. With these capabilities in mind, new network architecture is described below that enhances data credibility, quality, privacy and 'shareability', encouraging participation at personal, social and urban scales.

1.1 Opportunity

Researchers, policymakers and the public use data to understand and persuade; higher quality data tend to generate more significant action and better understanding. Los Angeles recently claimed the title of the Metropolitan Area Most Polluted by Year-Round Particle Pollution [1], so a close-to-home example of civic concern driven by data is the relationship between air quality and public health. This area is widely researched throughout the country using both top-down and bottom-up data collection, and suggests the potential for contribution by participatory sensing networks. [2] documents a community-university partnership that studied disproportionately heavy truck traffic through the Hunts Point area in New York City and its effect on local asthma rates. Local university researchers performed monitoring of particulate levels with specialized equipment and collaborated with the community to document commercial truck traffic, eventually linking traffic density with diesel exhaust particulate levels and uncovering an illegal use of non-designated truck routes—data that can influence public policy and health.

A new architecture for participatory sensing can enhance and systematize existing methodology by increasing the quantity, quality and credibility of community-gathered data. By implementing adaptive data collection protocols based on local or system-wide statistics, participatory sensing would facilitate accurate data entry by automatically geotagging, or enabling automatic upload from specialized equipment that is not yet network-connected. In professionally designed, top-down data collection projects like that in Hunts Point, data quantity could be increased by deploying prebuilt observation applications to existing mobile phones of willing community members—for example, for truck traffic counts. Data quality can be increased by implementing basic entry verification on the devices ("Did you really just see 40 diesel trucks go by in five minutes?") and by directing users to repeat missing or faulty measurements in real-time, before the environment changes. By tagging data points with *network-attested* location and time, a feature of the network

WSW'06 at SenSys '06, October 31, 2006, Boulder, Colorado, USA.

architecture introduced below, the credibility of information on when and where the data was taken is increased. Coverage can be enhanced and the community's time more effectively used by employing other adaptive protocols that prompt user entry based on location, time or knowledge of existing sample coverage and environmental factors. For example, a mobile application might request a community member spend his or her time counting diesel trucks in areas or at times of day where there is insufficient coverage in the existing data set.

With the use of mobile handsets, important new possibilities also emerge in the gathering of individual exposure and activity data. In addition to interactive data collection, periodic audio samples could be analyzed with simple environmental sound classifiers and automatically uploaded. For example, audio samples could be used to detect the probability of whether the device's user is in 'heavy automobile traffic' – a primary contributor to individual exposure to particulates. Additionally, mobile devices can be used to gather human activity patterns for future correlation with environmental data collected by the government and healthcare providers; such documentation of personal activity can aid doctors' analysis of individual patients exposure—as well as in analysis of community-wide activity and exposure data. The architecture outlined below provides mechanisms to enhance the privacy of those choosing to share such data.

1.2 'Grassroots' sensing

Participation means more than data collection: We believe that communities whose primary expertise (or concern) is not the enabling technology should be able to specify and execute data-gathering campaigns. With the right tools, professionals and community groups alike could employ participatory sensing campaigns to gather data about short-term concerns (e.g., changes in traffic or weather patterns, rerouted school buses with increased pollution exposure because they are now traveling on freeways) without waiting for a formal project or grant funding—yielding bottom-up, grassroots sensing. Citizens have intimate knowledge of patterns and anomalies in their communities and enabling them to respond is both empowering and valuable to long-term research. If such knowledge can be effectively gathered, it can also impact professional research and planning, as suggested by Jason Coburn in the *American Journal of Public Health* [3]:

"[I]ncreasing evidence in the natural sciences, public health, and urban planning reveals that expert assessments can miss important contextual information and need to be tempered by the experiences and knowledge offered by lay publics. Successfully reconnecting planning and public health will require the use of expert models, but it will also demand that these same models be recognized as contingent and fallible… Democratizing practice in both fields demands that professional knowledge not be compartmentalized from practical experience, that lay knowledge be considered alongside expert judgments, and that the incomplete models of the technically literate not be mistaken for the sum total of reality."

2. PARTISAN ARCHITECTURE

CENS researches and develops static and robotic mobile sensor networks. The greatest strides in these systems have been made through the use of multiscale data and models as context for *in situ* measurements and by employing in-network processing and mobility to achieve scalability in communication, energy, latency and coverage. We are now applying these lessons and technical approaches to participatory sensing systems. In contrast to the autonomous deployment approach of traditional sensor networks, *Partisan,* the architecture that we are developing to support this new use model, places users in the loop of the sensing process and aims to maximize the credibility of data they collect. We envision that Partisan will provide core network features including resource discovery, selective sharing, and location verification mechanisms, as well as application-level support for data gathering campaigns.

2.1 Core network services

Participatory sensing applications use data from a mobile sensor node gathered in collaboration with its owner/operator. Both explicit data gathering activities and background capture tasks are envisioned, with both being adjusted through local and network-provided feedback. They require fundamentally new algorithms and software mechanisms, because physical inputs and the actual location of fixed and moving devices become part of the critical data context. Network-level mechanisms are necessary to enable those who embed or carry sensors to share data in a controlled way while respecting the privacy of those being sensed. Additionally, the network should also provide basic quality checks for data and facilitate its discovery and dissemination within the privacy controls established. These fundamental network capabilities envisioned for Partisan can be summarized as context verification and resolution control; they are new network primitives intended to be implemented at the access point and router level, termed 'mediators' in Partisan. The architecture also includes application services for naming, dissemination and aggregation, which will be necessary to enable applications to take advantage of these core network features.

2.1.1 Network-attested context: location and time

Location and time are crucial context for mobile data gathering. In participatory sensing, they may be equally or more important to data credibility than the gatherer's identity. The more credible spacetime context is, the more useful the data for decision-making and the better 'grassroots' and widely-distributed mobile sensing works. In Partisan, the network itself *attests* to this context by tagging data packets with network-verified location and time, increasing data credibility and value. Essential to Partisan's building of spacetime semantics into the network fabric will be the ability to independently and verifiably timestamp and locate a device when it injects a packet into the network. Recent research [4-7] has explored verifiable location mechanisms that exploit physical and geometric constraints together with cryptographic mechanisms. For localization that uses the received RF signal strength to measure distance, work at UCLA [6] has shown that by utilizing hidden and mobile base stations together with a challenge-response mechanism and using the physics of signal propagation a network can achieve probabilistically verifiable localization to a given uncertainty.

2.1.2 Operating on physical context

The physical context of sensor data is richer than just the location and time that can be verified by the network. It includes, for example, the orientation of the sensor or node, measurement made by other sensing modalities, and measurements made by other sensors in the vicinity. Clearly, such additional physical context is of utility to subscribers in interpreting the sensor data or in checking its integrity. (Again, credibility of context is vital for participatory sensing, especially in the case of non-expert users.)

For example, the utility of sound level from a directional microphone is higher if the orientation of the microphone is also known. Sensor networks will increasingly employ ‘self-awareness’ [8] or ‘proprioceptive’ [9] sensors whose purpose is to acquire information about the physical context in addition to the directly measured phenomena of interest. Mobile handsets have many sensors that can be used for this during other data capture tasks. Partisan applies application-specific rules to calculate and verify the context. For example, an application may request that the network corroborate a sensor reading by comparing local readings with averages from nearby sensors. We envision that mediating nodes in our architecture (roughly in the same place in the network as current ‘access points’) will provide a suite of similar common aggregation functions over sensor readings. An application would then specify that the results from one of these functions over sensors in a geographical region be used as context. Mediating nodes may also maintain reputation information on their sensor sources, similar to that proposed in [10] for sensor data integrity and by [11] for Internet anomaly detection.

2.1.3 Context resolution control

Context increases the utility of sensor data. However, the publisher—especially one carrying a mobile device everywhere with them—may not wish to reveal too much. For example, the network may know a node’s location to a few meters but the subscriber may only be willing to share it to the zip code level. Likewise, a sensor may be willing to share information only as part of a geographical aggregate. Partisan is designed to encourage sharing by providing low-level mechanisms to control the resolution of data based on user-defined rules. Partisan’s mediators will be able to deliberately reduce the fidelity of the context information the network measures (location, time) or derives from sensor values. To combat emerging techniques for remote device fingerprinting based on measurements of timestamp drift [12] and localization using latency measurements [13], mediating nodes might add random jitter to packets. Besides physical context, network level identifiers such as host name or IP address also act as context. Our initial approach is to rely on the level of indirection provided by mediators to optionally hide a sensor’s network identity from subscribers. In case stronger anonymity is needed, a stream may be routed through multiple mediators, similar to how it is done for various anonymous routing schemes.

2.2 Discovery, naming and dissemination

Participatory sensing will require explicit, public and globally reachable naming, discovery, dissemination and aggregation services for data sources. To simultaneously protect privacy and encourage participation of the owners of those sources, Partisan’s approach to this challenge will implement selective sharing features for a hierarchical registry much like DNS. By enabling data name resolution requests to be dependent on the network-verified location and time of the requesting client as well as the selective sharing preferences of the data provider, a rich set of location-based privacy and anonymity rules can be built for discovery that complement resolution control of the data itself. Data aggregation will be implemented by higher-level services in the network but also named in the same registry as other data sources.

3. APPLICATIONS AND CAMPAIGNS

To guide development of the architecture, we have created a campaign model that can be used to formalize many participatory sensing applications, from geotagged photodocumentary to citizen-science data collection of public health factors. Rather than conceiving of data gathering in specific sensor deployments, we consider the campaign as a geographically and temporally constrained series of systematic operations to gather a particular type of data—using an already-deployed (but not at all static) network of mobile devices.

We can then define general user roles for such network-assisted, goal-oriented data gathering: (1) *Initiators*, who create campaigns and specify data collection challenges; (2) *Gatherers*, mobile users who participate in opportunistic data gathering that may be network-triggered, user-initiated or continuously captured, tagged and shared using the mechanisms described above; (3) *Evaluators*, who verify and classify collected data on behalf of the campaign; (4) *Analysts*, who process, interpret and present data and conclusions. The user types will have different perspectives on network function and specific services will be developed on top of the core Partisan features to support each of these roles. For example, gathering is fairly well supported directly by the network primitives described above. However, to allow initiators to communicate their sampling needs to gatherers (*i.e.*, the authoring of distributed data collection protocol and the triggering of opportunistic sampling based on that protocol) requires additional application-level services. We plan to develop methods for end users to define their own distributed data-gathering challenges, recruit other mobile users as participants, describe selective sharing rules and have tools to aggregate and analyze the set of observations made by the gatherers.

Four application areas have motivated our work and illustrate how this model might be used; they also reflect some of the many possibilities for participatory sensing.

3.1 Public health

Like the earlier air quality example, many possibilities exist for participatory sensing campaigns in public health: Individuals, health care providers, and community / government organizations could initiate opt-in activities to evaluate and support individualized and preventative care regimens, gather data for retrospective analysis of causes of chronic and environmentally-affected health issues, and generally to collect a wide range of high-fidelity health statistics for a population of interest. Autonomously captured and selectively shared activity pattern information could help chronic patients and their doctors link environmental factors with symptoms, while explicit data gathering might include automatic upload of at-home, self-administered diagnostic tests. Guided, geotagged data entry enables better compliance detection, timely trend and anomaly analysis, patient reminders and data quality feedback. Input mechanisms for data could include Bluetooth transfer of instrument readings when available; direct patient / caretaker input via structured text input in a local or web-based application; voice entry interpreted locally or on a server; input via cell phone-camera image of test output (display or strip) could be vetted by trained evaluators.

3.2 Urban planning

By lowering the complexity of creating trustable ad-hoc observing applications at the metropolitan scale, participatory sensing enables a very exciting application space for urban planning. Los Angeles is preparing for a two billion dollar redevelopment of a portion of its downtown, the Grand Avenue Project. The Norman Lear Center has invited citizen submission of design ideas for the project's park component, receiving and publishing hundreds of such submissions. Participatory sensing tools will enable these and other organizations to initiate data collection that similarly connect people (and their data) to the planning of their own environments. [14] describes a GIS-based noise planning tool created for the city of Belo Horizonte in Brazil, noting that noise is a major source of nuisance and, for many, an important quality of life metric. They model it in a GIS system but do not address how real world data might be gathered. A participatory sensing approach suggests that a simple service running on citizens' mobile devices, gathering and publishing basic statistics on ambient sound at regular intervals, with appropriate context checks, might be able to gather such data. Citizens could join a data-collection campaign to document noise levels in a community. They would configure simple selective sharing options to choose when and where samples are taken to calculate average sound amplitude, as well as the spatial and temporal resolution acceptable for network context tagging. A collaboratively generated city-scale analysis of noise levels at different times a day becomes feasible. When combined with participatory GIS techniques [15], incredible potential exists for developing important, accessible planning tools for communities of all sizes.

3.3 Cultural identity and creative expression

In 1996, Caroline Wang supplied women in a rural Chinese village with 35mm film cameras to document 'what is worth remembering and what needs to be changed'. They documented the results of a lack of adequate day care for children and midwifery training for women. As a result of presenting this work in a gallery seen by political leaders, local health policies for themselves and their children improved. [16] This became the Photovoice movement. Another set of motivating applications, new for sensor networks, seeks to combine the ethos of Wang's efforts with the increased ubiquity of image capture possible with network-connected, imager-equipped, always-on mobile devices. The decision to create a campaign to gather imagery might come from initiators within a community, and the Partisan architecture enables the network to lend some credibility to notions of when and where media is gathered. These features could allow the scale of participation to increase without losing the sense that the location was actually 'known' by the gatherer. Presence-based authentication (based on the time spent in a place rather than an identity record) is an important new possibility enabled by network-attested spacetime context. Paulos and Goodman propose that mobile technology should explore the role of 'familiar strangers', whose recurring physical proximity (but not interaction) add comfort to crowded city lives. [17] (See [18] and [19] for motivations in urban planning.) Partisan's mechanisms for recurring presence to be measurable by the network suggests that a distributed documentary vetted through 'presence-based authentication' is now possible.

3.4 Natural resource management

Finally, CENS is developing the ecoPDA, a mobile, manually-operated device designed to enable and increase the fidelity of field data gathered by environmental scientists and ecologists. Their paper-based data collection protocols have been translated into campaign descriptions, with field scientists acting as gatherers. Providing automated context tagging and uploading will minimize data gaps and data quality can be improved by providing some immediate feedback verification of correct data capture. In fact, both data quality and quantity are expected to improve over current manual and semi-automated models, including paper notebooks and direct PDA entry (without feedback). Image and acoustic sensing capabilities on the mobile devices can provide automatically analyed data streams and human interpretable data as well as quantify some subjective data inputs. They also can gather detailed, systematic metadata; image and sound annotations can be automatically linked to data entry so observation of anomalies can be retrospectively investigated. As primary data capture, media can be directly stored or converted either locally or on the server to numeric data using digital signal and image processing tailored to the specific context and tasks. As in in-home medical tests, visual test result indicators can be imaged next to calibrated reference charts, or images of foliage can be analyzed for density and other physiological characteristics as in remote sensing.

4. CONCLUSION

The personal mobile handset's penetration is unmatched in demographics, geographic coverage, acceptance and presence in everyday life. With image capture, positioning, connectivity and signal processing in new generations of devices, the basic components of a widespread participatory sensor network already exist. There is an exciting challenge to leverage the enormous public and private investment in wireless research and infrastructure to generate a proportional civic benefit. One area in which we see such promise is in participatory sensing as described above.

Effective participatory sensing will require more than ubiquitous mobile phones and 'mashable' web services. Above, we have outlined core network services and an application framework that we are using to develop applications that simultaneously protect privacy and encourage participation of handset owners. We believe these components, implemented on ubiquitous commodity hardware and overlaying existing network infrastructure, will enable what has been called 'citizen sensing', sensor network data blogging or 'slogging' [20], and allow our mobile devices and wireless infrastructure to act as resources for professionals and the public to gather vital information about the built and natural environment that was previously unobservable.

5. REFERENCES

[1] American Lung Association, State of the Air, 2004. http://www.lungusa.org/site/pp.asp?c=dvLUK9O0E&b=50752

[2] T. S. Lena, V. Ochieng, M. Carter, J. Holguín-Veras, and P. L. Kinney. Elemental Carbon and PM2.5 Levels in an Urban Community Heavily Impacted by Truck Traffic. Environmental Health Perspectives 110:10, October 2002.

[3] J. Corburn. Confronting the Challenges in Reconnecting Urban Planning and Public Health. American Journal of Public Health 94(4):541-549, 2004.

[4] S.Capkun, J.P. Hubaux, "Secure positioning of wireless devices with application to sensor networks," in Proceedings of IEEE INFOCOM 2005

[5] L. Lazos and R. Poovendran, SeRLoc: Robust Localization for Wireless Sensor Networks, ACM Trans. on Sensor Networks (TOSN), August 2005, Vol. 1, pp. 73 – 100.

[6] S. Capkun, J. P. Hubaux, ”Secure positioning in wireless networks,” IEEE J. Selected Areas in Communications: Special Issue on Security in Wireless Ad Hoc Networks, February 2006.

[7] S.Capkun, M. Cagalj, M. Srivastava, Securing Localization With Hidden and Mobile Base Stations, to appear in Proceedings of IEEE INFOCOM 2006.

[8] A. Kansal, J. Carwana, W.J. Kaiser and M.B. Srivastava, "Acquiring Medium Models for Sensing Performance Estimation," IEEE SECON, Sept 2005.

[9] J. Burke, J. Friedman, E. Mendelowitz, H. Park, M. B. Srivastava. “Embedding expression: Pervasive computing architecture for art and entertainment.” J. Pervasive and Mobile Computing 2(1):1-36, 2006.

[10] S. Ganeriwal and M. Srivastava, “Reputation-based Framework for High Integrity Sensor Networks,” Proc. 2004 ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN ’04), October 2004.

[11] Mark Allman, Ethan Blanton and Vern Paxson, "An Architecture for Developing Behavioral History," Proceedings of USENIX Workshop on Steps to Reducing Unwanted Traffic on the Internet, 2005.

[12] T. Kohno, Andre Broido, and K.C. Claffy, “Remote physical device fingerprinting,” IEEE Trans. on Dependable and Secure Computing, 2(2), 2005.

[13] S.M. Huffman, and M.H.Reifer, “Method for geolocating logical network addresses,” United States Patent 6947978, September 20, 2005.

[14] Oliveira, M. P. G., E. B. Medeiros, and C. A. Davis Jr. "Planning the Acoustic Urban Environment: a GIS-Centered Approach." ACM GIS '99, Kansas City, MO, 128-133, Nov. 1999.

[15] J. Corbett, et al. Overview: Mapping for change—the emergence of a new practice. Participatory learning and action 54:13-20, April 2006.

[16] Wang C, Burris MA, Xiang YP. Chinese Village Women as Visual Anthropologists: A Participatory Approach to Reaching Policymakers. Social Science and Medicine, 42 (10): 1391-1400, 1996.

[17] E. Paulos, E. Goodman. “The Familiar Stranger: Anxiety, Comfort, and Play in Public Places.” Proc. CHI 2004, Vienna, Austria. 2004.

[18] J. Jacobs. The Death and Life of Great American Cities. Vintage (reissue) 1991.

[19] M. Granovetter. The Strength of Weak Ties. In American Journal of Sociology (1973), vol. 78, p. 1360-1380.

[20] M. Hansen. ”Embedded.” Keynote, 2005 SIAM Intl. Con. on Data Mining, Newport Beach, California, April 21-23, 2005.