Urban Sensor Data Privacy Issues: Findings of the Array of Things (AoT) Privacy Breakout Group

Position paper for STREAM2015

Von Welch[1], Charlie Catlett[2]

Background: Array of Things Overview

The Array of Things[3] (AoT) is an urban sensing project, a network of hundreds of interactive, modular sensor boxes that will be installed around Chicago to collect real-time data on the city's environment, infrastructure, and activity for research and public use. This initiative has the potential to allow researchers, policymakers, developers, and residents to work together to evaluate and take specific actions that will make Chicago and other cities healthier, more efficient, and more livable.

The AoT nodes will initially measure temperature, humidity, barometric pressure, light, vibration, carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, ambient sound intensity, pedestrian and vehicle traffic, and surface temperature. Continued research and development will help create new techniques to monitor other urban factors of interest such as flooding and standing water. Additional sensors including particulate matter, precipitation, wind, carbon dioxide, methane, and other pollutants are being evaluated for inclusion in the sensor nodes. Because all of the data will be published openly and without charge, it will also support the development of innovative applications, such as a mobile application that allows a resident to track their exposure to certain air contaminants, or to navigate through the city based on avoiding urban heat islands, poor air quality, or excessive noise and congestion.

The AoT project includes a growing number of collaborators intending to deploy test configurations in nearly twenty other cities in the U.S. and globally. Additionally, the nodes will evolve over time with new sensors and processing capabilities. Thus a carefully developed privacy policy is essential and must both contemplate current capabilities and plans as well as governing future decisions about new features. At the AoT project kickoff meeting, an interdisciplinary group of university researchers, city policy makers, and private sector participants (see acknowledgements) met September 2-4, 2015 in Chicago IL. A breakout group met and discussed requirements around an AoT privacy policy. This paper summarizes that group's findings and may be guide other organizations dealing with streaming sensor data.

[1] Center for Applied Cybersecurity Research, Indiana University. vwelch@iu.edu

[2] Urban Center for Computation & Data, University of Chicago and Argonne National Laboratory. c@anl.gov

[3] arrayofthings.us

AoT Privacy Goals and Stakeholders

There are several goals for defining a set of privacy principles, and ultimately policy, for AoT: (1) Set responsibilities for system designers, developers, and operators; (2) define terms of use for data consumers; and (3) Inform populations of their rights and how the sensor network affects them. Different privacy stakeholders will be impacted by privacy responsibilities and protections in different ways:

- Data generators, the general population or “sensees” who are collectively the source of collected data;
- Researchers developing in situ data processing algorithms or other technology running on the sensor nodes;
- Data storers who archive or buffer any data;
- Data distributors, who provide the data to consumers;
- Data consumers who access and then utilize the data.

Particular Privacy Challenges of Urban Sensor Platforms

The privacy breakout group identified some particular privacy challenges of AoT urbana sensor data as compared to a standard IT systems. Three challenges identified as key include: (1) the Inability for an individual to easily opt in or out of participation in the AoT; (2) as a general-purpose, research-oriented, and continuously evolving platform, the AoT presents new challenges for privacy that aren’t present in previous sensor or technology deployments that are traditionally more task-oriented (i.e. bridge monitoring sensors); and (3) data collected or processed by the AoT could include measurements that are generally considered public (i.e. air quality) and those that are often considered private (i.e. images, albeit AoT is designed without the capacity to store images beyond what is necessary to process them locally on the nodes).

Privacy Breakout Group Findings

The break out group made some general findings around the AoT privacy policy, including:

- The privacy policy will define AoT’s balance of utility with potential harm to individuals or society. Each privacy decision is a decision balancing utility vs potential harm[4]. In the AoT, as is common with social or medical research, any future data measurements with potential privacy implications will be reviewed by University of Chicago's institutional review board (IRB) and an external, independent expert privacy committee.
- Understanding and communicating the utility of the data collection and sharing will help the population understand and accept said collection and sharing.
- There needs to be a code of conduct for researchers accessing in situ data on the sensors. While technical measures and reviews can, to a large degree, prevent them from accidentally or purposefully violating privacy, such techniques are not infallible and attempting to defeat them must be disallowed, with penalties defined.
- Data consumers have privacy rights and concerns regarding what data they access. There will be a tension around this as those funding the data distribution will often want

[4] E.g. http://sunlightfoundation.com/policy/opendatafaq/#safeguard

to know who is using it. This trade-off must be carefully considered. It was noted that there is resistance to requiring authentication or other barriers to data access that could be perceived as restricting access to the data.

- Key to the successfully definition and implementation of a privacy policy is the definition of a clear trust model between elements of the system - i.e. how each element is trusted to behave and how it can trust other elements.
- In situ analysis will be particularly challenging because of the potential use of algorithms or software from researchers outside the AoT project. Those researchers must clearly understand their responsibilities and their adherence will be validated and technically enforced.
- Architectural decisions can be made to eliminate some privacy risks. Some of these have already been made and should be codified as principles. For example, the architectural decision that images will be processed in real time and discarded, with no storage or transmission of images, is an implementation of a policy that the project has embraced and articulated, that “no data that could identify an individual will be transmitted from the nodes.” This policy should be formally captured in the privacy policy document.

Next Steps

The AoT is scheduled for deployment in phases beginning in the first quarter of 2016 with 50 nodes. A second phase with an additional 150 nodes will take place in late 2016, and a third phase adding 300 nodes will take place in mid 2017. The AoT privacy policy will be finalized and published for public comment prior to the first deployment.

Acknowledgments

At the Array of Things (AoT) Kickoff workshop, September 2-4, 2015, with 90 participants from Chicago and 14 other cities, comprising university researchers, city government representatives, and members of the private sector. Fourteen participants met to discuss the formulation of a privacy policy for the AoT project. Members of the breakout group included Brenna Berman (City of Chicago Department of Innovation and Technology), Janus Hoeks (Intemo), Bill Howe (U. Washington), Maggie King (U. Chicago), Lee W. Lerner (Georgia Tech), Lindsey-Paige McCloy (City of New York Mayor's Office of Technology and Innovation), Derek Meyer (U. Wisconsin), Michael Ruiz (Georgia Tech), Theo Tryfonas (U. of Bristol), Von Welch (Indiana U.), and Brant Zwiefel (Microsoft). The opinions expressed in this document represent personal opinions of some, and perhaps not all, members of the breakout group, and should not be interpreted as the position of any organization or project.