Overview

The vision of Participatory Sensing is of distributed data collection and analysis spanning the personal, neighborhood, and community scales. In Mobilize, student learners design, deploy, and analyze the results of data campaigns; while their fellow students act as data campaign participants, making key decisions about what, where and when to make observations using mobile devices to support the subject of the campaign. The Mobilize software system is designed to engage students in an integrated and iterative learning experience; one in which they choose the topic of the campaign, design the details of the experiment and participant experience, and most importantly, develop their computational thinking skills as they explore the datasets generated by their campaigns.

Approach

A data campaign is a coordinated data gathering exercise. A campaign consists of one or more surveys, each of which is a collection of prompted inputs, or questions. The answers to these questions provide insight into the subject of the campaign, organizing the capture of “facts” and personal observations from participants. To gather both widespread and timely responses, Mobilize distributes a custom mobile phone application for participants, ensuring that all responses are automatically timestamped, geocoded and uploaded for analysis and visualization.

The Mobilize software system supports expressive campaign “authoring,” mobile phone-based data capture and realtime feedback, privacy respecting data management, and extensible data exploration.

Campaign Authoring

A campaign author creates a configuration file (using the Mobilize XML specification) that describes the surveys, the survey content and the conditions around which they are delivered to participants. A campaign consists of one or more surveys. Each survey comprises of a set of questions/actions called prompts. Properties of individual prompts and their ordering are specified in the configuration file. After the campaign configuration is completed, it is then uploaded to the server and validated. A web application will allow students to publish their campaigns for use by other students and to manage active campaigns by displaying various participation statistics. These displays are themselves authorable by the students as they share a common language with the data analysis tool we will describe below.

Data Capture

A campaign participant downloads the Mobilize mobile application to their Android smartphone (in the future and contingent on available resources, additional smartphone types may be supported) and then downloads the XML for the particular campaign(s) in which they will participate. Participants can activate the application anytime to answer and submit survey questions. In addition to manual survey activation, a trigger mechanism is also provided. A trigger automatically launched a survey based on some certain criteria such as time and location. Mobilize currently supports relatively simple triggers, but these will be extended in future versions of the system. One specific class of triggers are exposed in the mobile phone application and allow participants to adjust their timing or frequency.

Feedback

Mobilize accommodates various forms of participant feedback. At a basic level these mechanisms are designed to sustain user engagement and participation. Our feedback schemes draw on data analysis metrics such as ranking among peers, campaign adherence, and progress towards goals. The first version of the system will have a very simple form of user feedback, however this function will be significantly enhanced in the second year of development.
Data management
The Mobilize data server maintains the data uploaded for each campaign in an SQL database. The database schema for a campaign is defined by the XML that is authored by the campaign owner, a student or teacher. Within a campaign, data are associated with the particular user that captured them. Moreover, when initially uploaded a participant’s data are not made accessible for a configurable period of time until the individual releases the data for sharing. This allows participants, who we consider to be the owners of their data, to decide to delete data that they do not want to share.

Data exploration
While the campaign is running and after the campaign participants have completed their activities, the student learner campaign authors will be able to download campaign data and will be provided with a rich, but scaffolded, set of tools with which to explore the data. During the course of the campaign, high level participation statistics will be generated automatically to support campaign management and engagement (such as number of responses, spatial coverage of the current responses, time-series analysis of data contribution levels, etc.), but all detailed exploration of data will be created by the student learners themselves, rather than be pre-scripted by the system. This design is essential to the pedagogical aims of Mobilize and will be explained further in the discussion of the curriculum.

Systems Description
The system consists of the following three components: Server, Web Interface, Mobile phone, and Data analysis support.

Server
The server component provides user authentication and authorization; supports upload, validation and storage of campaign configurations and survey responses and images; provides APIs for retrieving campaign configuration, survey responses and data points; and supports aggregate campaign data for feedback, data analysis and visualization.

The standard communication protocols, HTTP and TLS/SSL, are used to securely transport data between components. All text-based data (e.g., survey responses) is encoded in JSON format and send over HTTPS as content message. An image is sent as binary content over the HTTPS protocol. The server is designed to be as stateless as possible, so authentication information must be sent with each call to the server.

Two different authentication schemes are designed to authenticate clients. Thin-client authentication is designed for web browser clients. After a web user is authenticated using password, a temporary token is generated and sent to the browser to be stored as a secured cookie for later authentication. This token is also stored locally on the server. After tokens expire, log-in is required to generate new tokens for accessing the server. Thick-client authentication is designed for a thick client (e.g. phone application) that can keep local state. In this mechanism, a user’s password is securely stored locally on the phone. The same user’s credential is also created and stored on the server. This allows users to logon to a phone to interact with the application offline. The campaign configuration file is validated against the database schema. The campaign configuration is then used to validate incoming data uploads (e.g. survey responses) from the phone application.

These server functions are implemented in Java 6. They running on top of Tomcat 6.0, mySQL 5.1 database, and CentOS operating system.

Mobile Phone Application
The mobile phone application supports loading of campaign configurations; trigger functionality, including a GUI for users to adjust their trigger configuration such as changing trigger times, and the generation of notifications to remind participants to take actions; survey prompts based on menus, free text, and image input; buttons that allow users to initiate short surveys; mobility data capture in the background based on sampling of GPS, wifi, and accelerometer devices in the phone; and the upload of the captured data to the server. It also supports user login on the phone so that phones can be shared among data campaign participants.

The Android phone application is designed to read a campaign configuration and render appropriate prompt display according to the configuration. The campaign schema is shared between the phone and the server. A prompt type and properties associated with each prompt need to be specified in the configuration. This information tells the phone application on how to render and handle each prompt. Seven prompt types are currently supported.

Campaign configuration is an XML file. It contains properties and description of a campaign. A campaign configuration contains one or more surveys. A survey can contain prompts and repeatable sets of prompts with a rendering order defined by their top-down appearance within a survey and in conjunction with conditions. A prompt corresponds to a question or an action in a survey. Prompts should be designed to reflect facts, status, condition or observation of the object and its environment focused in the campaign study. A prompt configuration consists of 4 parts: prompt...
metadata (e.g., ID), web-based visualization properties (e.g., display type), phone-based prompt display properties, and branching condition (optional).

The phone application is implemented in Java and runs on the Android platform.

Data access, analysis and visualization
This component provides the key functions for student learners to access the data, and to support their development and exploration of computational thinking and analysis skills. It supports secure login, campaign management, data management, and basic campaign monitoring. Through a web interface authorized users download campaign data for local data analysis and insert additional web-based data visualizations that are created through their analysis. The analysis document can be uploaded to the system to be shared to other students.

The local analysis component consists of R and Deducer. R provides software environment for statistical computing, data analysis and graphics. On top of R, Deducer is used to provide an intuitive graphical user interface for students to perform basic data manipulation and analysis tasks without programming getting in their way. To fully support the Mobilize curriculum, Deducer will be extended and customized e.g., basic geo analysis, text mining and image analysis will be added.

Campaign analytics and data visualization aimed to help campaign authors monitor and run their campaigns are done at the server. We are using RApache, an Apache web server with embedded R interpreter, to perform actual statistical analysis and graphics. An RApache web application will handle campaign analysis requests through its interactions with the Mobilize server.

Accomplishments
Describe specific accomplishments (e.g., findings, deployments, system performance) during the reporting period.

Future Directions
Mobilize 1.0 will be released in June 2011 and used during our Professional Development and Summer Student Internship program during the summer of 2011. The initial version will have all of the functionality described, but in relatively simple forms. During the summer and fall of 2011 we will use our pilot usage experience to prioritize the expansion and deepening of features. Likely areas of focus for version 2 include: supporting richer forms of feedback and game mechanics in PS data campaigns, extending the support for continuous data streams such as activity (from GPS, wifi, and accelerometers on the phone), and communication (from audio and application statistics gathered on the phone), and adding tie in to social media applications. By pursuing our implementation through multiple cycles of iteration and use, we will prioritize our development to meet the needs of the Mobilize classrooms and learning objectives.

While the vision for Mobilize is that students will create their own campaigns, focused on issues of concern to them in their school and neighborhoods, we will seed this process by providing example canonical applications. These canonical applications serve three important purposes. First, it gives us concrete end to end tests of our functionality and performance during the months of software development, testing and integration. Second, it provides concrete areas in which to gather background data sets and organize specific lessons for the first year of the new curricular unit in ECS. And third, it provides a concrete experience for students and teachers to use the technology of Participatory Sensing, and based on this experience to begin to generate ideas for new and different campaigns.

We define more than one canonical application in order to test and demonstrate different aspects of the PS authoring, data capture and analysis system; and to provide a range of examples to students and teachers in order to communicate the breadth of potential applicability.