ECAR

An Introduction to BYOE Mobile Data Collection

April 2014

Rosemary A. Rocchio
Director of Educational and Collaborative Technologies, OIT
UCLA

An ECAR Mobile Strategy and Application Development Working Group Paper

Introduction

As of January 2014, smartphone ownership was at 58% among all Americans. However, the rate of ownership by college-aged Americans significantly surpasses this—83% for people aged 18–29.[1] The report *ECAR Study of Undergraduate Students and Information Technology, 2013* found that undergraduates own "two to three Internet-capable devices, and ownership of smartphones and tablets jumped the most (among all devices) from 2012 to 2013" and that students are "ready to use their mobile devices more for academics, and they look to institutions and instructors for opportunities and encouragement to do so."[2] As these numbers continue to increase and as students become more interested in using their devices to conduct coursework and research, campus IT organizations will need to support high expectations of ubiquitous wireless connectivity and to support students who want to use their devices for in-class activities and in-field research.

Institutions will also need to help students manipulate and store the data that are collected from these devices and assist the administration when it looks to students who are using these devices as a source of information. As students continue to arrive on campus expecting to leverage their devices to interact with their institution and their coursework through a broad array of mobile interfaces, higher education IT is presented with an opportunity to leverage the bring-your-own-everything (BYOE) trend to improve the campus environment and to experiment with new mobile data collection (MDC) research methods, which enable data gathering *in the field* using a mobile device.

Students have long used tablets and laptops to collect data in the field.[3] This type of MDC has typically been directed by instructors and researchers and conducted using university-owned

devices that are supplied to the students for that specific research purpose. The explosive growth of BYOE, however, has been a game changer, delivering in each new class of students an even more sophisticated "instrumented population." These mobilized "digital citizens" expect to use their devices to collect new data about their campuses and environments via crowdsourcing—where small amounts of data collected by many participants. Students have a relationship with their "companion" devices and are well versed in using them to measure their habits, publish their opinions, and track their activities for analysis.[4] Indeed, students will look to solve issues in innovative ways, leveraging both public and, when available, private crowdsourcing methods.

In this context, MDC is the practice of gathering data in the field via a *personal* mobile device. BYOE has made data gathering simultaneously cheaper and easier. At the same time, however, the emergence and ubiquity of smart mobile devices that run different device-specific (native) apps that are integrated with the phones' higher-level data-collection instruments (e.g., the camera, audio recorder, accelerometer, and GPS) make using BYOE for MDC complex. Collecting data in the cloud or through a public data aggregator (such as Twitter) is one way around these issues. IT organizations that can provide a platform for their campus to conduct private MDC will allow their students to become active researchers, gathering field data that can then be examined and visualized in aggregate in their courses and research projects for deeper and more engaged learning.

Mobile data collection holds promise while also raising concerns that range from pedagogy to privacy, security, and other policy and research implications. These questions are still emerging in this nascent area of research, however. This paper serves as an introduction to the four main types of MDC—passive, public, crowdsourced, and private—highlighted with exploratory use cases that demonstrate how learning, research, and community engagement can leverage this new technology space.

Passive Mobile Data Collection

Passive MDC takes place when a user's device collects data streams without any overt interaction by the user other than installing an app—for instance, data from the device's accelerometer, GPS, SMS log, or call log. Data collected passively may later be combined with active data streams, such as survey data, for deeper analysis.

Passive MDC Example: Actigraphy

Actigraphy—in which data about a person's rest and activity cycles are collected, often using a custom device such as a pedometer—can be combined with passive data collected from a mobile phone or similar device. For instance, combining actigraph data with location data from a

phone's GPS allows map visualizations of the data. Figure 1 is a visualization of a passive mobile data collection on the day that the collector ran the LA Marathon. Collecting such data can be used in higher education to gauge the use of bike paths on campus, in health studies, and more.[5]

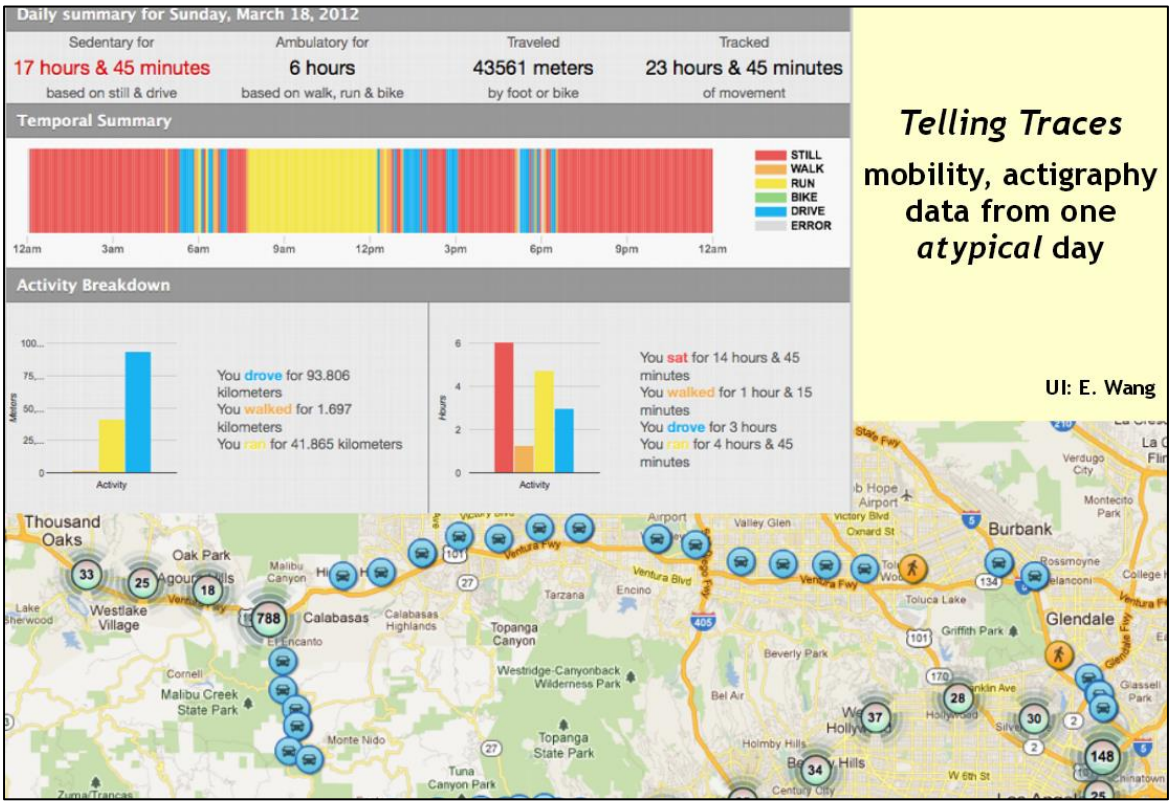

Figure 1. Visualization of Passive Mobile Data Collection
http://www.cs.cornell.edu/~destrin/resources/talks/2012-july-mobilizinghealth.pdf

Public Mobile Data Collection

Public MDC leverages data from the plethora of social media apps—such as Twitter, Facebook, and Foursquare—and the fact that it has become popular for users worldwide to create public data postings that can then be culled by researchers. Very interesting visualizations can be created when these data are combined with location information, serving as rich fodder for understanding a variety of events in recent history, from the Arab Spring to the Occupy Wall Street movement to the relief efforts in Japan after the Fukushima nuclear incident. In the case of public MDC, the research intent is defined *after* the public data have been created, allowing researchers to take advantage of existing data sets. However, researchers need to act fast, as many public data feeds are available for only limited periods of time.

Public MDC Example: Twitter Aggregation with GIS

One example of public MDC combines location information with Twitter postings following the Japan Tohoku earthquake and tsunami (see figure 2). Volunteers from several institutions worked to cull this information using various hashtags, including those tagged with pleas for help, for information about evacuation, and more.

Figure 2. Japan Tohoku Earthquake
http://gis.ats.ucla.edu/japan/

Public MDC Example: Social Media and GPS Data Visualization

University administrations can also take advantage of MDC and students' use of BYOE. One such example is Social.Chapman, an effort to collect content from Chapman University's social media community—from platforms such as Facebook, Twitter, Instagram, and WordPress—and aggregate it on the university website in real time.[6] Collecting data in real time lets the Chapman community know "what is happening now." In addition, this public MDC allows the institution to maintain a searchable, archived database of the content from campus social media accounts to help the institution identify university trends, evaluate branding strategies, and monitor the consistency of the community's response.

In addition, the school has pulled this information into what they call Inside Chapman, where—with the addition of Google Analytics (web analysis and reporting), SharedCount (a URL tracking platform), Disqus (a commenting platform), and a unique time-degradation formula—the school community can discover content based on popularity, sort stories based on "what's trending," and integrate campus blog posts (in addition to public social media data).

Crowdsourced Mobile Data Collection

Crowdsourced MDC is the practice of using the general public or an online community—sometimes referred to as "citizen scientists"—to contribute collected content without compensation. The concept of using crowdsourced data has grown in popularity recently as a solution to common challenges. For instance, researchers in astronomy have leveraged crowdsourcing techniques to classify galactic images.[7] In fact, the use of this type of MDC in higher education is also growing, not only to solve issues on campus (such as where more trash receptacles are needed)[8] but also to enable academic field research that would otherwise not be possible.

Crowdsourced MDC Example: Tracking Invasive Plants

Initiated by UCLA and now maintained by the University of Georgia, the What's Invasive! project takes advantage of citizen scientists to supply data (see figure 3). In this case, data are collected from localized contributors (e.g., hikers) willing to help by downloading an Android app, taking photos of invasive plants, and submitting them via the app, along with GPS information. With the data collected, scientists are able to locate these invasive plants and analyze growth patterns. As of March 2014, with just over 300 people participating, the project had gathered nearly 11,000 observations and identified over 200 invasive species in 113 active sites across the United States and even farther afield, in Scandinavia, the Middle East, and Asia.

Figure 3. What's Invasive! Community Data Collection
http://www.whatsinvasive.org/index.cfm

Private Mobile Data Collection

Whereas public MDC takes existing publicly posted data and applies analysis to that data, private MDC gives the power of defining the mobile data survey instruments to faculty, researchers, and students so they can identify what specific data to collect. The data (both passive and intentional) are then collected from a defined group of participants using mobile devices and uploaded to a server for private aggregation, research, and analysis. Both the collection process and the data collected are private in this type of MDC. Image, GIS, timestamp, video, audio, and accelerometer data lends itself to this kind of collection, though many other forms of data could be gathered in this way. Data sets gathered through private MDC can be collected for research (keeping in mind that if using human subjects the study will require IRB approval), class field study, self-study, community improvement, health monitoring, political opinion gathering, and more.

In addition to using data gathered from BYOE devices, private MDC sometimes uses data from private social networks. An increasing number of higher education efforts create private social networks—using tools such as Yammer, Jive, and Chatter—to help increase collaboration and encourage data sharing within a university community. As these private social networks grow, we can expect to see an increase in private MDC efforts taking advantage of the data they have to offer.

Private MDC Example: STEM Education in High Schools

The Mobilize Project is an instructional and research program from UCLA's Center for Embedded Network Sensing and Center X, the NSF, the Computer Science Teachers Association, Google, and the Los Angeles Unified School District (see figure 4).[9] Based on the open-source platform ohmage,[10] the project is an effort to engage high school students with STEM and computational thinking by creating "Participatory Sensing hands-on, inquiry-based curricular units and teacher professional development for computer science, mathematics, and science high school classes."[11]

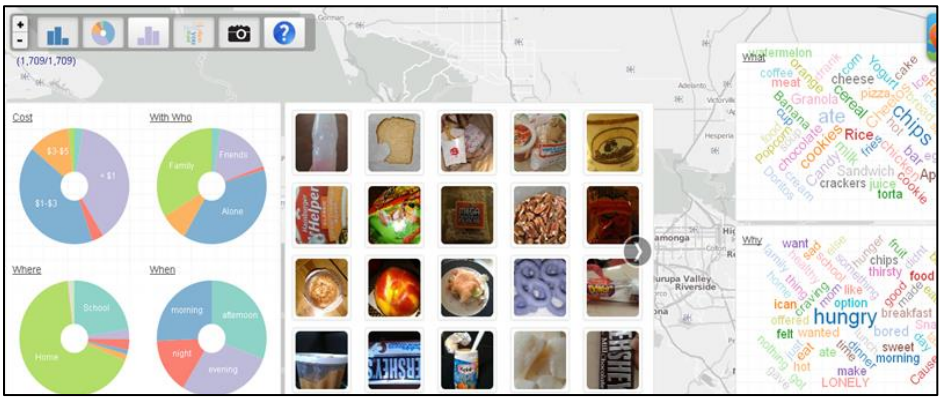

Figure 4. Mobilize: Snack Data Visualization
https://lausd.mobilizingcs.org/snackdemo/

Private MDC Example: In-Field Classwork

In the winter quarter of 2013, a UCLA labor relations studies class used the multiplatform ohmage tool to collect data from labor union members across campus. The class worked together to define the survey instrument they would later use to conduct interviews; they would use mobile devices to upload the data at the conclusion of each encounter. Their final survey instrument included 33 questions and took about 10–15 minutes to complete with an interviewee. The class was able to conduct 225 interviews across five teams. The data that each team collected was analyzed, visualized across each team, and combined with data from the other teams for further aggregate analysis.

Conclusion

Mobile data collection is a growing trend in researchers' toolkit, one that incoming student populations are eager to embrace. Courses that involve students' learning basic research principles and activities (developing hypotheses and then collecting evidence) are likely to be early adopters of MDC platforms. Researchers with a desire to collect new data about study participants, such as location-triggered surveys, will also be looking for MDC solutions. Private MDC open-source platforms (such as ohmage) do exist and are expected to grow in popularity.

Providing faculty, researchers, and students with enabling technologies like BYOE MDC platforms is an exciting opportunity for campus IT departments. It exemplifies what it means for our organizations to embrace the BYOE trend and give our campus communities supported methods to experiment more, with dimished risk, less cost, and greater engagement.

Contributors

Special thanks go to the following ECAR Mobile Strategy and Application Development (ECAR-MSAD) Working Group contributors to this report. If you are interested in the topic of mobile web and applications for higher education, visit http://www.educause.edu/ecar/ecar-working-groups/mobile-strategy-and-application-development-working-group.

Mary Ann Blair
Director of Information Security
Carnegie Mellon University

Serge J. Goldstein
Associate CIO and Director, Academic Services
Princeton University

Mark Katsouros
Director, Network Planning and Integration
The Pennsylvania State University

Linh Pham
Senior Systems Analyst
University of Southern California

Brett Pollak
Director, Campus Web Office
University of California, San Diego

Lori Tirpak
Director, Enterprise Systems
Oakland University

Notes

1. Pew Research Center, Internet & American Life Project, "Mobile Technology Fact Sheet," http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/.
2. Eden Dahlstrom, J. D. Walker, and Charles Dziuban, *ECAR Study of Undergraduate Students and Information Technology, 2013*, research report (Louisville, CO: ECAR, 2013), 5, https://net.educause.edu/ir/library/pdf/ERS1302/ERS1302.pdf.
3. See, for example, Meg Stewart, Jeffrey J. Clark, Jeremy William Donald, and Keri Van Camp, "The Educational Potential of Mobile Computing in the Field," *EDUCAUSE Quarterly* (March 29, 2011), http://www.educause.edu/ero/article/educational-potential-mobile-computing-field.
4. Many students arrive on campus already using their mobile devices for "self-monitoring," but this new trend can be given a broader name: the "Quantified Self" movement (see http://en.wikipedia.org/wiki/Quantified_Self). In this movement, people actively seek to acquire data on a variety of aspects of their lives, including inputs (e.g., food, vitamins, and medicines consumed) and performance output (miles/steps jogged or walked, weights lifted, etc.), sometimes using automated sensors such as GPS data and locations logged and at other times using self-reporting for more subjective measures such as mood and feelings. Adding to this trend is the ability to wear video cameras on helmets (cyclists and skiers/boarders), in watches, and in glasses.
5. See, for example, John Hicks, Nithya Ramanathan, Donnie Kim, Mohamad Monibi, Joshua Selsky, Mark Hansen, and Deborah Estrin, "AndWellness: An Open Mobile System for Activity and Experience Sampling" (proceedings of Wireless Health 2010: Academic and Research Conference, La Jolla, California, October 5–7, 2010), http://research.cens.ucla.edu/people/estrin/resources/conferences/2010-oct-wireless-health.pdf. The article discusses a platform developed at UCLA for the "real-time assessment of a participants [sic] health and behavior," and describes "an initial qualitative study plus several planned future studies" that use the platform.
6. For more information about how social sharing data are being used to rank Chapman content, see http://blogs.chapman.edu/smc/2013/10/03/ranking-content-with-social-metrics/.
7. See, for example, John Matson's blog post, "Crowdsourcing the Cosmos: Amateurs Sift Through Astronomical Data," *Scientific American*, February 18, 2009, www.scientificamerican.com/blog/post/crowdsourcing-the-cosmos-amateurs-s-2009-02-18/. In this post, he describes how 150,000 volunteers helped submit classifications on over a million images.
8. Deborah Estrin and Jeff Burke, "URB 11 Participatory Campaigns for Sustainability (GarbageWatch), Citizen Science (Networked Naturalist), and Active Living (CycleSense)," Center for Embedded Network Sensing, 2009 Annual Report, http://research.cens.ucla.edu/urban/2009/urban11.pdf.
9. Mobilize: Mobilizing for Innovative Computer Science Teaching and Learning, http://mobilizingcs.org.
10. For more on ohmage, an "open-source participatory sensing technology platform," see http://ohmage.org.
11. Mobilize, "What is Mobilize?" http://www.mobilizingcs.org/about.