The emergence of smartphones as a viable sensing platform has radically shifted the landscape of mobile-sensing research. Mainstream consumer smartphones equipped with a range of sensors are the everyday computing platform for millions of people worldwide. This has dramatically lowered the barrier to building and deploying large-scale sensing systems that can operate not only on a personal level but also at a community scale, sensing entire cities and beyond.\(^1\)

Even years after early stage prototypes,\(^2,3\) researchers are still converging over how to design smartphone sensing systems. A key open question is precisely how to leverage these systems’ ability to sense beyond a single individual as they collect and interpret sensor data. In this direction, we’ve made steady progress in our ability to automatically extract and reason about complex social structure and dynamics using data readily available to smartphone systems. For instance, macroscale mobility patterns and social network topologies can be mined from the Call Detail Records (CDRs) of mobile phones.\(^4\) At a more personal scale, coarse proximity data from Bluetooth radios can distinguish different types of social ties,\(^5\) or fine-grain characterizations of face-to-face interactions\(^6\) can be inferred from wearable sensors’ audio streams.

Yet the design of existing smartphone sensing systems still focuses most strongly on the individual, carefully considering issues such as usability or user engagement. This is entirely appropriate given the mobile phone platform’s highly personal nature, but such considerations also imply that people live in a vacuum, unaffected by others’ behavior. In fact, individuals are tightly connected via various social and physical processes. Phenomena including social influence,\(^7\) homophily (the tendency of similar people to form social bonds),\(^8\) and information diffusion\(^9\) collectively cause various types of interpersonal dependencies. Users form complex hierarchies of densely connected community groups and are influenced by group behavior and social networks. By neglecting the relationships and networks that link users, we’re ignoring an important part of the design space for smartphone sensing.

What if sensing systems comprehensively understood the complex community dynamics within a user population? How should they be designed and function given such awareness? This column’s underlying conjecture is that we can revolutionize smartphone sensing if it becomes community-aware, with systems empowered to see beyond just a single user and designed to understand and react to a set of rich social phenomena.
and fluctuating connections within the wider community.\textsuperscript{10}

To make this discussion more concrete, I look at three mobile-sensing scenarios. For each one, I consider how community-aware sensing can alter existing practice. Specifically, I examine how a community-aware approach can lead to radically new ways of sensing communities in a distributed manner, allow individual users to assume more sophisticated roles within these systems, and enable sensor-based models of human activity to generalize to large user populations despite the presence of significant inter-user diversity.

**Mobile Health**

Using mobile phones to continuously monitor fine-grained physical activity\textsuperscript{11} or a broader set of health outcomes could change how society diagnoses and treats medical conditions. Such mobile systems are rapidly advancing in the areas of sensing, inference, persuasion, and resource management. However, virtually all of them assess health using data from individual users. This ignores an important dimension of personal health: how friends, family, and others influence users’ health and well-being. For example, monitoring exposure to communicable illnesses — critical to at-risk people such as the elderly or those with certain chronic conditions — has limited effectiveness if we consider only the individual. However, if systems understand the mobility patterns and activities of both individuals and those they routinely interact with, we can combine such information with epidemic models that capture the aggregate implications of social interactions. This will better equip sensing systems to track the risk factors of developing diseases (especially in relation to particular behaviors). In support of this direction, researchers have recently demonstrated the effectiveness of monitoring changes in socialization, such as face-to-face interactions, to act as an indicator for contagious diseases such as common colds and influenza.\textsuperscript{12}

Similarly, personal interactions can significantly influence health-related behavior, such as an individual’s eating habits\textsuperscript{13} or mental state\textsuperscript{14} (see Figure 1). A system that understands the dynamics of such interactions — and the overall social network in which they occur — can have a much broader perspective on such aspects of well-being. By combining microscale, on-body sensor data with macroscale awareness of external social factors, mobile health systems will be better equipped to not only capture but also predict health outcomes.

**Crowdsourcing and Human Computation**

Exploiting “people-in-the-loop” within mobile phone sensing systems is a natural direction given that a smartphone is a personal device with which users frequently interact. Earlier work has studied the potential for leveraging crowdsourcing to intelligently collect mobile sensor data.\textsuperscript{15} Now, hybrid systems are emerging that blend the strength of human computation\textsuperscript{16} with machines to produce systems with otherwise unachievable operating points (such as trade-offs between accuracy and delay).\textsuperscript{17} However, the conventional view of these hybrid systems maintains a narrow perspective on people — that is, the systems are designed so that the same type of loop is used repeatedly for everyone and only one person is embedded in it.

Crowdsourcing achieves scale by creating millions of these individual loops. The approach, while undeniably powerful, misses crucial opportunities by treating everyone as homogeneous cogs in a machine. In reality, a group’s suitability for performing crowdsourcing tasks (for example, collecting or processing data) varies from group to group, depending on the skills and attributes of the individuals involved. What’s missing are automated methods to

---

**Figure 1.** The correlation between the social connections (graph edges) among people (nodes) and a categorization of daily food choices (node size) for a study conducted in Beijing, China. More generally, substantial evidence from medical and social science exists linking health outcomes to a variety of social phenomena. However, virtually none of the existing mobile health systems consider such connections when assessing users’ health.
Beyond Wires

Figure 2. The Community Similarity Networks (CSN) framework. The framework trains personalized activity classifiers sensitive to the interpersonal differences and similarities among users. CSN incorporates various similarity dimensions (such as shared physical and lifestyle traits) enabling activity recognition to better generalize to diversity in the user population.\(^{19}\)

identify and characterize user communities. Once available, this information can be used to tune a system’s operation and more easily maintain design requirements (such as accuracy and cost).

For instance, for a crowdsourcing system to effectively perform a data interpretation task, users in the community must share characteristics such as location, knowledge, skills, and available free time. Other sensing systems require just the opposite and need diversity. One example is crowdsourced experiments in which users are the study participants (such as a user study), which normally require forming diverse groups representative of an entire population. Through careful, automated selection of user communities that perform those sensor-based crowdsourcing tasks (such as data collection or interpretation) they’re best suited for, people-in-the-loop systems can achieve higher quality and less variable results. Crowdsourcing systems unaware of a community’s structure will be subject to unpredictable performance that depends on the composition of a group of users who happen to accept an offered bundle of system tasks. Existing work in the context of matching individuals to roles within participatory sensing campaigns\(^{18}\) can likely assist in this area.

Activity Recognition

In this final example, I consider a recently published activity-recognition framework, Community Similarity Networks (CSN),\(^{19}\) because it adopts community-aware principles to overcome challenges to large-scale mobile classification.

As a user population’s size increases, so does the amount of diversity it contains. Users can vary for several reasons, a clear example being physical differences such as height, weight, or gender. People can live and work in different places and have different cultures and socioeconomic origins. Although they might do the same basic set of activities (workout, go to work, socialize) they can do these activities in very different ways. Mobile systems commonly use various supervised classifiers (example-based) to recognize behavior and context. However, these classifiers fail to generalize to such levels of diversity — leaving accuracy unpredictable and dependent on precisely who happens to use the system.\(^{10,19}\)

CSN approaches this problem by using different dimensions of interuser similarity (that is, physical, lifestyle, and sensor-data-driven) to build a weighted user graph, a similarity network that captures the various ways users are alike (see Figure 2). Under CSN, each similarity network is embedded within a chain of classifier training algorithms to emphasize labeled training data between users who share common traits. As a result, CSN can personalize generic classifiers to each user’s different characteristics. Using a variety of similarity networks is necessary because no single type of similarity is effective across all the categories of activities. For example, physical characteristics are useful for activities such as exercise, running, or climbing stairs. Lifestyle characteristics (such as occupation, temporal patterns, and where people live and work) are more beneficial to activities such as transportation-mode inference.
Community-Aware Smartphone Sensing Systems

CSN clearly demonstrates how community awareness built into sensing systems at design time allows naturally occurring phenomena, such as communities with shared behavioral patterns, to overcome challenging sensing problems.

**Toward Community Awareness for Smartphone Sensing**

Community-aware smartphone sensing systems will need a deep understanding of the diverse ties that connect people and cause networks of communities to form within diverse user populations. Such systems will leverage this understanding while performing basic sensing system operations to fundamentally change how these operations occur. Noticeably, the three examples provided all use communities in very different ways. At this nascent stage of community-aware sensing, we aren’t yet close to a generalized framework for leveraging communities within smartphone systems. The complexities and sheer variety of community interactions and social ties within large-scale populations are extensive, leading us to ask, which of these interactions and ties are the most salient to the operation of sensing systems? Can we embed sufficient understanding of these effects into system components without resorting to developing systems on a case-by-case basis?

Central to community-aware systems’ evolution will be the development of **community models** — computational models of the networks that bind individuals — that can mine the various links between individuals, recognize different community types comprising tightly connected people, and determine these communities’ collective characteristics. Figure 3 illustrates two key roles these models will have within community-aware smartphone sensing systems — specifically, influencing how phenomena are sensed and guiding a variety of system operations (such as data collection) and optimizations. An expanded understanding of community dynamics will alter how these systems sense real-world phenomena ranging from traffic patterns to user preferences toward places of interest\(^{20}\) or even smartphone application usage.\(^{21}\) Community models will play a role in sensing whenever connections between people can affect the phenomena in question.

Similarly, we can embed community models into key sensing operations to guide how these components function or optimize their performance. A clear illustration of this is CSN, in which the entire workflow of training activity-recognition models changes to incorporate the similarity networks that exist in the user population. Another example is the sensing systems that leverage crowdsourcing discussed previously; in such systems, community models can intelligently optimize the selection and assignment of communities to specific sensor data collection or interpretation tasks. By introducing the community model and leveraging its aggregate predictive power, we can enable sensing systems to

---

**Figure 3. A community-aware smartphone sensing system. Central within these systems will be the community model, which captures the social structure within a user population and the communities that form. The model can both impact how the system senses phenomena and guide how various sensing system operations are performed and optimized.**
optimize common system trade-offs, such as cost, accuracy, energy, and latency, despite unpredictable user behavior.

To progress toward general-purpose community-aware smartphone sensing systems, we’ll need more sophisticated community models that will in turn demand new approaches to sensing at both the micro-(personal) and macro- (community) scale. Even more pressing is the need to find ways to seamlessly connect these two viewpoints in the design and architecture of large mobile-sensing systems, which would let us migrate from existing examples of community-aware sensing that fail to generalize beyond single narrow domains. Finally, privacy remains perhaps the most critical obstacle.

The extensive awareness of communities envisioned, even if it were possible today, requires information that the general public would be uncomfortable sharing. We’ll need methods for either performing this type of analysis using privacy-preserving techniques or developing systems that provide suitable transparency (such as audit trails) that we can balance against people’s reservations.

Acknowledgments

Thanks to Andrew Campbell (Dartmouth College), Tanzeem Choudhury (Cornell University), Feng Zhao (Microsoft Research Asia), and Lin Zhou (Shanghai Jiaotong University) for their valuable feedback.

References


Nicholas D. Lane is a researcher at Microsoft Research Asia. He’s an experimentalist who likes to build prototype smartphone sensing systems based on well-founded computational models. Lane has a PhD in computer science from Dartmouth College. His dissertation pioneered community-guided techniques for learning human behavior models better able to cope with real-world conditions containing diverse user populations. Contact him at niclane@microsoft.com.