Privacy and Security in Mobile Health: A Research Agenda

David Kotz, Dartmouth College
Carl A. Gunter, University of Illinois at Urbana–Champaign
Santosh Kumar, University of Memphis
Jonathan P. Weiner, Johns Hopkins University

Mobile health technology has great potential to increase healthcare quality, expand access to services, reduce costs, and improve personal wellness and public health. However, mHealth also raises significant privacy and security challenges.

With the advent of miniaturized sensors, low-power body-area wireless networks, and pervasive smartphones, the burgeoning field of mobile health (mHealth) technology has attracted tremendous commercial activity, consumer interest, and adoption by major healthcare providers. This technology has great potential to increase healthcare quality, expand access to services, reduce costs, and improve personal wellness and public health. These benefits will only be achieved, however, if individuals are confident in the privacy of their health-related information and if providers are confident in the security and integrity of the data collected.

The US spends more than $2.6 trillion annually on healthcare. This amount represents approximately 18 percent of the country’s gross domestic product, a percentage that has doubled over the past 30 years and is the highest of any nation in the world. Over 75 percent of these costs are due to the management of chronic diseases, which currently affect 45 percent of the US population. By 2023, the annual costs to manage chronic diseases alone are expected to rise to $4.2 trillion. Similar challenges occur in many developed nations with an aging citizenry, and in developing nations that strive to provide better healthcare to their growing populations.

Numerous countries look to IT—increasingly, to mobile technology like smartphones and wearable sensors—to address these problems. However, health IT faces broad software assurance challenges, and overcoming these will be critical to adopting mHealth systems and realizing their benefits. Privacy and security were cited as the most important concerns in a recent survey of 27 “key informants” from across the US healthcare and mHealth sectors. Furthermore, a year-long Washington Post study of cybersecurity revealed that “healthcare is among the most vulnerable industries in the country, in part because it lags behind in addressing known problems.” The recent breaches of
Health-insurance giants Anthem and Premera underscore this point. Here, we focus on the privacy and security challenges of mHealth technology.

**Health IT Privacy and Security Challenges**

Health IT systems face daunting security and privacy challenges due to six recent trends:

- The locus of care is shifting as the healthcare system seeks more efficient and less expensive ways to care for patients, particularly outpatients with chronic conditions.
- Strong economic incentives to keep patient populations healthy, rather than caring for patients only when ill, are motivating healthcare providers to pursue innovative prevention plans and treatments of chronic conditions that entail more continuous patient monitoring outside of the clinical setting.
- Mobile consumer devices like smartphones and tablets are quickly being adopted by patients, caregivers, and healthcare providers for health and wellness applications in addition to their many other uses, making it difficult to protect sensitive health-related data and functions from the risks posed by general-purpose devices connected to the Internet.
- Significant emerging threats target health IT systems, while new regulations strive to protect medical integrity and patient privacy.
- Rapid technology advances that enhance mobile devices’ utility—for example, computational models that convert wearable-sensor data into measures of addictive behaviors such as cocaine use or smoking—increase the range of potentially private events that can be inferred from seemingly innocuous sensor data.
- Healthcare organizations lack the technology and expertise to adequately secure patient data; according to a recent survey, 69 percent of clinicians said their organization did not address demonstrated cyber vulnerabilities in medical devices approved by the US Food and Drug Administration (FDA).

These trends are driving major changes in the health IT landscape, and require research to develop effective security technologies that work across care settings and support continuous data collection in the context of multi-purpose mobile devices.

Before exploring the challenges in detail, we first define our scope. Traditional approaches to securing healthcare systems have relied on isolation, using tools like firewalls and network access control. However, the trends...
SECURITY THREATS

described above make it unfeasible to simply “lock down” medical devices or health-records systems, especially because patients and staff use part of the system outside the clinical context and many of the wellness applications of this technology are entirely non-clinical. Instead, these trends demand “wide-spectrum” security technologies that can be adjusted to fit the system user’s needs and expertise. A major healthcare provider has professional staff that can configure and monitor security settings in its electronic medical record (EMR) database, but an individual patient must have intuitive and hassle-free security technologies for home-based devices.

Given this scope—mHealth technology used by individuals who might be supported by caregivers and providers, perhaps remotely—we can specify numerous open research challenges that span technology, policy, and organizational domains.

DATA SHARING AND CONSENT MANAGEMENT

Most mHealth systems collect data about a person’s physiology, physical activity, or social behavior and are designed to store the data for later analysis by caregivers and providers. Data sharing raises the question of consent: how and when does the person decide whether, and with whom, to share what data and at what level of granularity?

In the traditional health information management model, patients consent to the collection and use of their personal health information (PHI) for treatment purposes. Further consent is often sought for additional PHI uses, such as research.

mHealth systems, however, often collect a far broader range of information, much more continuously and for a wider range of uses than is collected in traditional clinical settings. Research is needed to help individuals understand what data is being collected, where it is stored, who has access to which data at what granularity, and what it will be used for. Indeed, individuals should be given personal preferences regarding PHI collection, dissemination, and retention. Regulations such as HIPAA (the Health Insurance Portability and Accountability Act) and HITECH (Health Information Technology for Economic and Clinical Health) in the US provide some guidance but do not apply to much of the personal wellness domain, and leave wide latitude for creative abstractions and interfaces that would allow people to make informed choices about their PHI.

Research challenge: how can an mHealth system expose to its users, in an understandable way, what data is being collected, what information is being shared with whom, what might be inferred by that information, and where and how the information might be used, and then notify users of any deviations from the agreed-upon protocol?

ACCESS CONTROL AND AUTHENTICATION

User consent or policy determines who can access mHealth data, but how do mHealth systems confidently identify the individual(s) they are sensing or who is using the system? Identification is critical to attach the correct identity to the mHealth data for provenance, and authentication is the foundation of access control and audit logging.

Many of today’s mHealth apps are based on a smartphone, leveraging its sensors and user interface to collect, process, and report health-related information about the device’s owner. As smartphones are designed as personal devices, it is normally safe to assume that the user is indeed the owner. Of course, a smartphone can be stolen or borrowed by another person, resulting in the phone’s mHealth apps recording data about the wrong person to the owner’s health record or exposing the owner’s PHI via app displays and notifications. It is thus important for a smartphone to know when it is not in the owner’s possession. Most work on this problem focuses on initial authentication to unlock the user interface (most commonly via numeric codes, swipe patterns, or fingerprints), but there is a real need for continuous authentication—that is, repeatedly verifying that the phone’s holder is the person who initially authenticated.

Many future mHealth apps will use wearable devices to measure activity, behavior, and physiology and even to directly influence the body. Such devices must be able to verify the wearer’s identity to ensure that the collected data is posted to the correct health record and that any treatment applied is truly intended for the wearer. One solution is to build biometric sensing into the device, such as the bioimpedance approach taken by Cory Cornelius and his colleagues.9

Furthermore, any method for identifying and authenticating smartphone or wearable device users for mHealth apps must be accurate, applicable to most persons, robust to environmental conditions, unobtrusive, and resistant to various attacks.

Research challenge: develop continuous user authentication
methods for mobile devices such as smartphones and smartwatches that suspend data collection, personal notifications, and access to personal data when the device is used by someone other than its owner.

CONFIDENTIALITY AND ANONYMITY

Much of the information—whether physiological, behavioral, or social—collected by mHealth systems is sensitive and highly personal. The data must remain confidential, subject to access-control policies and mechanisms, and anonymous when used for research and public-health purposes where individual identities are not necessary.

Anonymization

Mobile-sensor data provides researchers with unprecedented opportunities to quantify the complex temporal dynamics of key physical, biological, behavioral, psychological, social, and environmental factors that contribute to disease. For example, GPS data makes it possible to collect geo-exposures (such as proximity to a tobacco point of sale for a newly abstinent smoker or to a fast-food restaurant for a congestive heart-failure patient) and movement patterns (such as driving or physical activity), and to study their impact on health.

However, mobile-sensor data can also disclose private information about the user. For example, GPS data can reveal not only the user’s identity but also all the places the user has visited, some of which might be private. Even if GPS is turned off, data collected by the accelerometers and gyroscopes embedded in smartphones and smartwatches for activity monitoring could be used to characterize a person’s movement patterns.

Sharing raw mobile-sensor data thus carries re-identification risks. Sharing only high-level inferences—for example, begin/end times at home or work—from the data might limit such risks but also significantly limits the data’s utility.

Research challenge: understand and quantify re-identification risks inherent in various mobile sensors, and develop data-transformation methods to limit such risks while retaining scientific utility.

Behavioral privacy

Measurements from mobile devices and wearable sensors can provide unique visibility into a user’s health status, stress, addictive behavior, eating patterns, sedentary behavior, geo-exposures, and daily social interactions. Such data can help researchers better understand the etiology of complex human diseases responsible for more than half of all US deaths. However, sharing this data also poses new privacy challenges. For example, audio data can reveal conversational and emotional characteristics, exposure to TV programming and advertisements, and video game playing and other activities, but it can also capture private and intimate details.

There is a need for technologies that mitigate the risks of behavioral privacy disclosure while also supporting the health or wellness goals for which the data is collected. For example, real-time audio processing could be used to extract relevant health inferences while discarding sensitive content but would necessitate improved algorithms. Likewise, breathing patterns could be used to infer conversation episodes but would require wearing respiration sensors and would not capture either conversational content or speakers’ identities.

Research challenge: understand and characterize privacy disclosure tradeoffs inherent in sharing behavioral data, and develop data-transformation methods to limit privacy risks while retaining the data’s scientific utility.

Continuous and unintended sensing

Continuous long-term data streams from various sensors entangle useful health-related data with information about user identity and behaviors. For example, reviewing audio recordings of conversations with one’s spouse (perhaps in conjunction with a therapist) could help improve marital life, but continuous audio recordings can also capture private conversations with nonconsenting persons, which is unethical and in some jurisdictions illegal. Requiring users to manually turn sensors on and off is burdensome as well as prone to frequent compliance failures.

Research challenge: develop mechanisms that can automatically turn sensors on and off to preserve user privacy and can be personalized to minimize user burden while maximizing utility.

Multiplexed sensor semantics

A key benefit of mHealth sensors is that the same sensor can be used to infer various behaviors. For example, electrocardiography can be used to monitor cardiovascular health, but ECG can also be used to infer stress level and the use of some drugs, such
SECURITY THREATS

as cocaine. Similarly, smartwatches can capture activity levels but can also infer eating and smoking behaviors from hand gestures. Inferring behaviors and health states from sensors is a rapidly evolving field; each new research result increases both the utility of an existing sensor and its inherent privacy risks. Hence, characterizing the behavioral information content of a specific sensor is difficult.

Research challenge: create computational mechanisms to ensure that users can control the inferences made by an authorized entity receiving sensor data streams.

mHEALTH SMARTPHONE APPS

Many mHealth benefits will be delivered to users, caregivers, and providers through smartphone apps. These apps might

– use the phone’s sensors to record sounds, take photos, or record motion;
– communicate with other sensor devices worn on the skin or collect health-related information from nearby sensors that, for example, sense contaminants in the air; or
– collect data from the user’s EMR in a hospital or from a cloud repository.

This wide range of possibilities has aroused concerns about the techniques used to secure mobile devices and mHealth apps. Much of the smartphone app market lies outside government regulation, although the FDA and Federal Trade Commission have started to address these concerns in the US. The quality of implemented security measures varies widely.11 Some recommendations are available for mHealth app developers, and mobile device management (MDM) solutions can help clinical enterprises secure smartphones and tablets. There is also a promising proposal to develop a “building code” for safety-critical medical systems.3

Research challenge: develop best practices for securing mobile devices and their apps, and develop platforms that will provide these benefits at low cost.

Current smartphone app architectures also raise privacy concerns. In particular, the Android platform, which makes up 80 percent of the smartphone OS market, has a degree of openness that supports strong innovation but also puts users at risk of privacy violations. These concerns arise from two aspects of the Android architecture. First, the degree of information flow between apps is worrisome because the wide range of apps likely to populate the average user’s smartphone creates a possibility that at least one app will gather information about other apps on the device and use it in ways the user might not approve of. Second, apps commonly incorporate advertising libraries, which means they effectively share their privileges with advertisers, weakening the “least privilege” principle and opening the threat of privacy leakage via advertising libraries.12

Research challenge: clarify threats to, and develop security and privacy protections for, smartphone apps that handle medical and health data—in particular, develop methods to isolate apps from advertisers.

POLICIES AND COMPLIANCE

Access to mHealth systems and the information they provide is typically managed by policies, which might emanate from consumers (as when they indicate data-use preferences), the operating procedures of healthcare providers or technology organizations, or government regulations. Policy development and enforcement results from a complex interplay of multiple stakeholders. Because technology is essential to help monitor and enforce these policies, policymakers must understand the wide and evolving range of relevant technologies.

Research challenge: What technical mechanisms could enforce data-management policies as mHealth data is collected, stored, processed, and shared? Could technologies developed for digital rights management (DRM) assist in ensuring that an individual’s personal privacy preferences remain attached to data about them, and that these preferences are enforced even as the data is stored and forwarded to providers and other health-care system participants?

To realize the promise of mHealth devices and applications, everyone involved—from patients to providers to payers—must trust the system to provide high-integrity data and services while respecting users’ privacy. This trust is partly based on mechanisms built into the technology, including cryptographic protections on data at rest and in transit, access-control policies, and authentication mechanisms. Ultimately, though, trust resides in the people and organizations manufacturing and distributing
devices, developing software, operating services, and using the data. The trust relationships among these actors, and the legal and regulatory frameworks that support those relationships, are a critical foundation for the technological mechanisms.

Research challenge: What ecosystem supports mHealth? Who are the stakeholders, and what are their roles? What policy and legal frameworks need to be in place for them to serve these roles? What standards need to be developed, and what certification mechanisms can encourage and ensure compliance with the standards?

Thus, there is a need to map out a conceptual trust architecture for mHealth systems that identifies the

› various types of actors;
› natural trust relationships, such as between patient and physician; and
› legal frameworks—for example, contractual relationships between a healthcare provider and a cloud provider, or a regulatory relationship between a government agency and device manufacturers.

A conceptual mapping would provide a clean abstraction for reasoning about the security and privacy properties of mHealth systems, and could guide creation of a regulatory framework in the real world. The World Health Organization recently reviewed key aspects of the current state of this regulatory framework across the globe. The framework, while progressing rapidly, is still in the earliest stages of development in most nations.

Research challenge: determine the most effective way to help develop, manage, monitor, and enforce consumer-directed, organizational, and government policies and regulations associated with data collection and use within the mHealth ecosystem.

ACCURACY AND DATA PROVENANCE

For mHealth systems to achieve their full potential—improving healthcare, reducing costs, and expanding access—those receiving information produced by these systems must be able to trust their accuracy and veracity.

In addition to the threats posed by common cyberattacks, the physical coupling of sensors and actuators make them vulnerable to attacks mounted from the physical channel, such as signal manipulation. To protect not only data but system inferences and decisions, solutions to such attacks must go beyond traditional cryptographic mechanisms and employ novel techniques from control theory, game theory, and other disciplines.

In our conversations with physicians and researchers, one of the most frequently cited concerns about mHealth data collected outside the clinical setting relates to the data’s authenticity and accuracy. The data must be tagged with information about the data’s provenance—what device collected the data and what was done to the data—as well as the context in which it was collected. This metadata must be securely bound to the data with a combination of cryptographic hashes and signatures to ensure that neither the data nor metadata has been tampered with.

Such methods might be feasible in simple situations where a sensing device is uploading raw data directly to the recipient’s health-data server. In many advanced applications, however, the data passes through multiple stages of processing including filtering, summarization, aggregation, and combination with other data sources. What is the best way to convey information about all these data sources and processing steps?

Contextual information is even more difficult to define and collect because it often depends on the type of health data being collected. For a blood-pressure reading, for example, it is important to know whether the subject applied the cuff correctly to her arm, rested her arm on a flat surface, and remained still throughout the reading. Aarathi Prasad and her colleagues proposed one approach to the specification and collection of contextual evidence for mHealth sensor data, but much more needs to be done to recognize the many factors that affect the quality of such data.

Research challenge: develop extensible methods for collecting, storing, and presenting contextual information along with health-related data collected by mHealth devices and apps to help data consumers verify and interpret the health data.

SECURITY TECHNOLOGY

Ultimately, many mHealth security and privacy approaches will rest on technological foundations; ideally, digital electronics for mHealth devices and apps will be designed with security and privacy in mind. Specifically, there is a need to

› identify hardware and software enhancements that would
help enforce users’ privacy preferences;
› protect the contents of mobile and wearable devices including PHI, cryptographic keys, and software;
› preserve the privacy of user context—location, device presence, communication, activity, and so on;
› create a secure execution space on mobile devices for handling health-related data;
› allow multiple software and services to coexist on mobile devices, without conflict, to enable software updates to be securely installed; and
› easily manage user authentication, data collection, and manageability—for example, remote disable and remote updates.

Of course, any solution to these problems must consider device resource constraints such as memory, CPU speed, bandwidth, battery life, and the user interface.

Research challenge: How should mobile-device hardware and software architecture change to help inform and protect individual privacy—specifically, to secure critical computations and data, securely store cryptographic secrets, and identify and authenticate the user?

Homomorphic encryption enables cloud-based servers to store and process sensitive mHealth data without those servers or their operators ever handling the unencrypted information, allowing mobile and wearable device users to leverage the power of cloud computing without needing to trust cloud services with this confidential data.16

Many mHealth technologies produce a large, long-term stream of data about a person’s health and health-related behaviors that, if aggregated, presents a huge opportunity for public health research. Imagine, for example, the potential benefits of tracking a million-subject cohort for a decade or longer, as envisioned by President Obama’s Precision Medicine Initiative (www.nih.gov/precision-medicine-initiative-cohort-program). The challenge is providing researchers with scientifically robust data from such a dataset without exposing individuals’ private information. Emerging differential privacy methods have great promise to achieve this dual vision.17

Research challenge: develop efficient homomorphic encryption techniques for mHealth data, and limit the amount of noise that must be added to data to satisfy differential privacy requirements.

The increasing capability and decreasing size of mobile technology offers many opportunities to improve health and wellness. The same technology, however, could cause users harm if the hardware and software systems are not designed with security and privacy in mind. The research community has an important role to play in developing effective, efficient, and usable mechanisms to secure mHealth technology and protect users’ PHI. To that end, we encourage our colleagues to address the many research challenges outlined in this article.

ACKNOWLEDGMENTS
The authors are supported by the National Science Foundation under award numbers CNS-1329686, CNS-1330491, CNS-1212901, and CNS-1213140; by the Department of Health and Human Services (SHARP program) under award number 90TR0003-01; and by the National Institutes of Health (NIH) under grant U54EB020404 awarded by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) through funds provided by the trans-NIH Big Data to Knowledge (BD2K) initiative.
The views and conclusions contained in this article are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors.

REFERENCES
7. B. Eastwood, “Premera Says Data Breach May Affect 11 Million

ABOUT THE AUTHORS

DAVID KOTZ is the Champion International Professor in the Department of Computer Science at Dartmouth College and serves on the US Healthcare IT Policy Committee. His research interests include security and privacy, pervasive computing for healthcare, and wireless networks. Kotz received a PhD in computer science from Duke University. He is an IEEE Fellow, a Senior Member of ACM, and an elected member of Phi Beta Kappa. Contact him at kotz@cs.dartmouth.edu.

CARL A. GUNTER is a professor in the Department of Computer Science and the College of Medicine at the University of Illinois at Urbana–Champaign, where he is also director of the Illinois Security Lab (ISL) and the Health Information Technology Center (HITC). His research interests include programming languages, formal methods, and security and privacy for healthcare and the power grid. Gunter received a PhD in mathematics from the University of Wisconsin. Contact him at cgunter@illinois.edu.

SANTOSH KUMAR is a professor and the Lillian and Morrie Moss Chair of Excellence in Computer Science at the University of Memphis. He is also director of the National Institutes of Health (NIH) Center of Excellence for Mobile Sensor Data-to-Knowledge (MD2K), which gathers investigators in computing, engineering, statistics, medicine, and behavioral science from 12 universities. His research interests include the practical and theoretical aspects of mHealth wireless sensor networks. Kumar received a PhD in computer science and engineering from the Ohio State University. He is a Senior Member of IEEE and ACM. Contact him at skumar4@memphis.edu.

JONATHAN P. WEINER is a professor of health policy and management and health informatics, and director of the Center for Population Health Information Technology (CPHIT) at the Johns Hopkins University’s Bloomberg School of Public Health and School of Medicine. His research interests include the application of electronic health records and health IT for population-based applications within communities and integrated delivery systems. Weiner received a DrPH in health services research from the Johns Hopkins University. He is codeveloper of the Johns Hopkins Adjusted Clinical Groups (ACG) System, case-mix/predictive modeling software used across the world to help manage the care of more than 100 million patients. Weiner is also a Fellow of Academy-Health and a member of the American Public Health Association and the American Medical Informatics Association. Contact him at jweiner1@jhu.edu.


