Adaptive Mobile Sensing: Leveraging Machine Learning for Efficient Human Behavior Modeling

(Technical Paper)

What Sensitive Data-Sharing Examples Tell Us About the Potential for a National Health Monitoring System in America

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Hannah Katinas
Fall 2019

Technical Project Team Members
Erin Barret
Cameron Fard
Charles Moens
Lauren Perry
Blake Ruddy
Shalin Shah
Ian Tucker
Tucker Wilson

Approved by Prof. Laura Barnes, Department of Systems Engineering

Approved by Prof. Kent Wayland, Department of Engineering & Society

General Research Problem

Is a national health monitoring system feasible?

In recent years, technology has become so intertwined in people's lives, that millions of Americans have become addicted to the information that these devices can provide, such as fitness analytics or sleep trackers (Sapacz, Rockman, and Clark, 2016). Coupled with a continuously advancing healthcare field, it is possible for healthcare organizations to rely on these devices to collect data and monitor patient activity for their benefit, but with increasing data privacy regulations, it is understandable why the healthcare field is so cautious. There is a smaller-scale system that is currently being studied regarding students living on college campuses, which is the focus of my technical paper. Through an application on participant's phones, we are able to collect various types of data that allow us to have a better understanding of their movements and activity throughout the day. This will hopefully allow us to detect an illness or injury before the participant detects it themselves. The STS research portion will focus on the feasibility of expanding this to a national health system by looking at other examples of similar systems that have been implemented. My technical team is already struggling with privacy issues and how to anonymize the information that we are collecting, so it led me to think about if this would even be possible at a national level if there are already privacy concerns with paid participants. With the research I will perform, I hope to explore the existence of a perfect balance between leveraging the power of smart devices while also avoiding any privacy concerns from participants.

Reliable Analytics for Disease Indicators: Leveraging Smart Devices to Predict Health

Which data collection algorithm will limit battery usage of activity monitoring applications?

Today, smartphones and other wearable devices are capable of collecting millions of data points about each of its users daily. However, while the potential power of this data in improving society and providing other benefits is unprecedented, there is still much work to be done in creating predictive models that can extract valuable information from this data. In the Reliable Analytics for Disease Prediction capstone project, such unstructured smartphone data will be analyzed in an effort to create predictive health models.

The technical project, advised by Professor Laura Barnes, Mehdi Boukhechba and Lihua (Lee) Cai, specifically seeks to predict the user's health status based on smartphone-extracted contextual data. The project is a part of ongoing research conducted for the Defense Advanced Research Projects Agency (DARPA) to design and develop reliable disease detection analytics through data collected from smartphones. The ultimate goal of the research is to create "a mobile application that passively assesses a warfighter's readiness immediately and over time," (Patel, n.d., para. 5); by building predictive health analytics that utilize smartphone sensors, the onset of illnesses, concussions, or even mental health issues can be noticed in real-time. In the current stage of research, the technical team will develop the tradeoff between data collection frequency and battery life. This is an important step in the feasibility of this technology and in understanding the user's environment. By understanding these limitations, accurate predictive models can be built without the obstruction of dead phones or other unwarranted stimuli.

Mobile sensing data used in this research will be collected through the Sensus App. This app, developed at the University of Virginia (UVA), uses "event-driven architecture that triggers actions in response to changes to the device or network state" (Lockheed Martin & Advanced Technology Laboratories, 2017, p.10). This data will be utilized to create context recognition models, which determine what ambulatory state the user is in, like walking, running, or sitting. Additionally, the Sensus app will push surveys as notifications to the participant's mobile phones to create additional context around the data collected. These surveys will include questions about the user's activities immediately before opening the survey, such as the user's location, length of activity, phone position, and more. This additional collected data will allow the team to build a strong base case for these predictive health models.

The technical project group consists of nine undergraduate Systems Engineering students. Because of the large size, the team is divided into three subteams: the Data Modeling Team, the Data Visualization Team, and the Data Collection Team. These teams were constructed for the currents needs of the project, and are subject to change and overlap depending on the need in each area. The Data Modeling Team will work to prove the efficacy of adaptive sensing in an attempt to find a balance between data collection and battery usage. Ultimately, the team will develop an algorithm as a potential alternative to the adaptive sensing model currently being used. The Data Visualization Team will make significant improvements to the web-based visualization platform used by the researchers to increase the understanding and context of the data they are collecting. Improvements to this platform will allow better insights to be easily accessible. The Data Collection Team is tasked to complete an Institutional Review Board Protocol, which is a document outlining how we will be gathering and securing the data that we will be using so that the data collection among the student cohort can begin. Once the IRB is

completed and approved, the team will be responsible for organizing the participants in the study.

At the end of the study, the team will deliver a recommendation for smartphone data collection that effectively accounts for a user's battery life and critical predictive data, as well as a recommendation for intuitive data visualizations for the researchers' web platform. The technical project is funded through a grant provided by DARPA. Additional resources include test phones and desktop computers to run software and view data. The technical project will produce a conference paper for the Systems Information Engineering Design Symposium (SIEDS) that will take place in May 2020.

Feasibility of American Trust and Willing Participation in a National Health Monitoring System

What sensitive data-sharing examples tell us about the potential for a national health monitoring system in America?

Introduction and Background

In America, the healthcare system is beginning to have the technology necessary to advance national health care by having several components that are required for early diagnosis and early prevention. (Kupwade and Seshadri, 20). The potential benefits of a national system that monitors its citizens' health data include faster illness detection in individuals, a decrease in dangerous mental health effects, lower health insurance rates, and decreased medical costs (Elbogen and Johnson, 2009). The only reason this is not an immediate goal for the near future is that there are so many privacy concerns regarding American citizens handing over their sensitive

information to any government or corporation. With smaller-scale examples of similar systems, I am hoping to find situations in which participants in the system willingly handed private data over to an organization. One example of a smaller-scale system like this is the Swedish national health care system, which uses a unique PIN number for all permanent residence in Sweden (Ludvigsson, Otterblad-Olausson, Pettersson, & Ekbom, 2009). Even though the residences do not give their explicit approval for this system, the national health care organization believes that it benefits not only the individuals, but also society as a whole. Looking at these different cases and drawing conclusions from them will be much easier and less costly than performing experiments or implementing this system on a number of willing participants.

Evidence and Data Collection

In this STS research paper, I will draw some similarities between this potential health monitoring system and existing systems from around the world. There are several key points in the search for evidence and cases that I will be focusing on in order to find examples that effectively relate to this potential health monitoring system. The first element that I am interested in investigating is cases with similar privacy concerns in the health field. Since the introduction of online databases storing patient health information, health concerns have increased dramatically due to misuse of data (Xu, 2019). Looking at health-related privacy concerns, I am interested in how the organization was able to find a way around the privacy concerns, or if they implemented data security programs or similar systems that eased the participants' concerns. The other element that I will be focusing on in evidence and sources will be the idea of people willingly, or unknowingly, giving their sensitive data over to an organization. This can either mean people deliberately taking action to provide their personal data or that they are unaware

that an action they are making has the effect of them giving out their personal data (Fazlioglu). For example, if someone was to download an application that claimed to use GPS location to find nearby stores, but the application actually used it to pinpoint stores they went into for more personalized advertisements, how would this affect their willingness to allow the application to use the GPS for its original purpose? These systems that I am interested in focusing on will mostly be found in published journal articles that have data and statistical findings to support their conclusions. With such a new area of study, I am expecting to find situations in which systems like the national one I had introduced earlier were attempted on a much smaller scale. But with enough systems to examine, I think drawing conclusions regarding the feasibility of a national health monitoring system will become easier.

Data Analysis

With the key elements that I am looking for in evidence described above, the overall goal is to draw on points of similarity between each system that I find, and the proposed national health monitoring system. One of the sources I have found discusses the privacy concerns and the high security regarding the telemedicine field, which is healthcare services provided over the phone or internet (Zaidan, Zaidan, and Kiah, 2011). It explores the impact of data privacy and confidentiality on telemedicine applications, which involves health-related information that participants are not usually willing to give out over the internet. More specifically, it discusses why participants are hesitant when it comes to these new technologies in the health field, and what the telemedicine field has done to combat these concerns. With more sources like this, I will first find meaningful similarities between the system at hand and the potential national health system. Then I will discuss how the steps taken to decrease the privacy concerns in the

existing system can be transferred over to the national health monitoring system. Additionally, I plan to focus on aspects that may include, but are not limited to, which groups are for or against the new system, what their concerns are, if the benefits seem to outweigh the resources put into the new system, and what constraints there are against the new system that are preventing it from evolving further.

Conclusion

With the benefits that come with a national health monitoring system, there are also obvious concerns regarding data privacy. By the end of my STS research project, I will have a better understanding of the privacy concerns that most people have when it comes to sensitive health data, as well as what is currently being done to counter these hesitations. This understanding of these concerns, paired with examples of similar existing systems, will help me determine what the national health monitoring system needs to do/have in order to be considered a feasible option. Through my technical research, my team and I will create an algorithm for collecting data from participants in our study that maximizes the context of the data, while minimizing the battery consumption of the Sensus application. Within my thesis, I look forward to exploring not only the feasibility of the physical tracking system, but also people's willingness to hand over private health data.

References

Elbogen, E. B., & Johnson, S. C. (2009). The Intricate Link Between Violence and Mental Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, *66*(2), 152. https://doi.org/10.1001/archgenpsychiatry.2008.537

Fazlioglu, M. (2019) Beyond The "Nature" of Data: Obstacles to Protecting Sensitive Information in the European Union and The United States. 37. *Fordham Urban Law Journal*, 271-306

Kupwade Patil, H., & Seshadri, R. (2014). Big Data Security and Privacy Issues in Healthcare. *2014 IEEE International Congress on Big Data*, 762–765. https://doi.org/10.1109/BigData.Congress.2014.112

Lockheed Martin & Advanced Technology Laboratories (2017). DARPA warfighter analytics using smartphones for health (WASH) ReADI technical section. (Grant proposal) Cherry Hill, NJ: Lockheed Martin and Advanced Technology Laboratories

Ludvigsson, J. F., Otterblad-Olausson, P., Pettersson, B. U., & Ekbom, A. (2009). The Swedish personal identity number: Possibilities and pitfalls in healthcare and medical research. *European Journal of Epidemiology*, *24*(11), 659–667. https://doi.org/10.1007/s10654-009-9350-y

Patel, T. (n.d.). Warfighter analytics using smartphones for health. Arlington, VA: DARPA: Defense Advanced Research Projects Agency website: https://www.darpa.mil/program/warfighter-analytics-using-smartphones-for-health

Sapacz, M., Rockman, G., & Clark, J. (2016). Are we addicted to our cell phones? *Computers in Human Behavior*, *57*, 153–159. https://doi.org/10.1016/j.chb.2015.12.004

Xu, Z. (2019). An empirical study of patients' privacy concerns for health informatics as a service. *Technological Forecasting & Social Change*, *143*, 297–306. https://doi.org/10.1016/j.techfore.2019.01.018

Zaidan, B., Zaidan, A., & Kiah, M. M. (2001). Impact of Data Privacy and Confidentiality on Developing Telemedicine Applications: A Review Participants Opinion and Expert Concerns. *International Journal of Pharmacology,* 7(3), 382-387. doi: 10.3923/ijp.2011.382.387