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

(Technical Paper)

Social Challenges of Uber's Driver Rating System

(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

Ian Tucker
Spring, 2020

Technical Project Team Members
Erin K. Barrett
Cameron M. Fard
Hannah N. Katinas
Charles V. Moens
Lauren E. Perry
Blake E. Ruddy
Shalin D. Shah
Tucker J. Wilson

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Introduction

As modern society moves further into a digitalized era, there comes an increasing struggle to balance consumer data collection and data privacy. Some journalists, such as Lawrence Summers from the Washington Post argue that data collection provides an ultimate good for society, such as providing predictive information on disasters or enabling accountability of policymakers (Summers, 2016). However, journalists such as Tim Sparapani from Forbes magazine argue that data collection creates an unfounded breach in individual privacy, which can pose a threat, particularly if illegal characters such as thieves and human traffickers gain access to databases (Sparapani, 2019). Regardless of such sentiments, there continues to be a push for developing new data collection and analysis methods that can maintain the privacy of users, even when the data is used in practical applications.

Such practical applications are known to heavily rely on either wearable technology (such as smartwatches) or, as is the case for this research, smartphones. The result is rapid drain of the user's smartphone battery life, which further disincentivizes users to allow data collection. Thus, the purpose of the proposed technical research is to develop varying adaptive models for data collection that will adjust to the user's behavioral patterns involving their smartphone, collecting data only when necessary to obtain meaningful information. The model should decrease the battery load on the user's smartphone whilst maintaining quality of data. The ethical effects of such data collection on society will be explored in the proposed STS portion of this research. In particular, the STS topic will explore the case of Uber, and how its driver rating system affects different social groups under Uber's labor market, namely full-time immigrant employees often marginalized by the ratings. Understanding of this system will inform developers on metrics and social factors to consider when aggregating data into a single driver rating.

Technical Topic

Today, smartphones and other wearable devices are capable of collecting millions of data 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 efficiently extract valuable information from this data. In the Reliable Analytics for Disease Prediction capstone project, such unstructured smartphone data will be analyzed as part of an effort to create predictive health models.

The technical project, advised by Professor Laura Barnes, Medhi Boukhechba and Lihua (Lee) Cai, specifically seeks to predict the user's health status based on smartphone-extracted contextual data. The project is 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.); by building predictive health analytics that utilize smartphone sensors, the onset of illnesses, concussions, or even mental health issues will 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 consumption. Said development is an important step in the feasibility of this technology and in understanding the user's environment. By gaining a better sense of these limitations, accurate predictive models can be built without the noise of dead phones or other unwarranted stimuli.

Mobile sensing data used in this research will be collected through the Sensus Application. 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 and 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 that ask questions about the user's activities immediately before answering 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 the strong foundational truth for these predictive health models.

The technical project group consists of nine undergraduate Systems Engineering students and has been subdivided into three subteams: the Data Modeling Team, the Data Visualization Team, and the Data Collection Team. 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 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 designated to complete the IRB 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 and a recommendation for intuitive data visualizations for the researchers' web platform. The technical project will produce a conference paper for the Systems Information Engineering Design Symposium (SIEDS) that will take place in May, 2020.

STS Topic

Rating systems are a common phenomenon used to assess the performance or quality of a person or item. However, one class of rating systems, social rating systems, which rate people on a multitude of factors relating to their performance in a certain action area, have become to center for some debate. From financial credit ratings like Fair Isaac Corporation (FICO) scores to student grading systems in public schools, social ratings systems are not unfamiliar to the public. With protests from drivers against the Uber driver rating system (Dickey, 2014) and the advent of China's social credit system (Philipp, 2018), a discussion on how certain social groups are disenfranchised more than others from such systems has emerged.

Uber provides an interesting case study due to its relative youth and its driver rating system detailing very clear consequences for drivers with low ratings; that is, *risk of deactivation* if a driver's rating falls below 4.6 stars (Cook, 2015). With different segments of Uber's driver market carrying various amounts of autonomy in relation to Uber, different social groups receive varying levels of pressure from the driver rating system, particularly as it pertains to opposing bias from riders (Rogers, 2015). For instance, part-time drivers, who usually consist of well-off non-immigrant drivers, tend to enjoy much more flexibility in their work and less stress from ratings (Bowman, 2019) and less bias to work against for their ratings. Meanwhile, full-time drivers, who are typically immigrants or minorities (Hua & Ray, 2018) must work hard to dissipate preconceived notions of themselves from riders to obtain higher ratings (Rogers, 2015). The research and analysis for the STS topic will focus on how social groups influenced Uber's driver rating system and generated such a divide.

Given that this topic seeks to explore the relationship between social groups (in particular, driver workforce segments in Uber) and the driver rating system, an STS theory will explain the driver behind this relationship. In particular, Social Construction of Technology (SCOT) will be used to explain the relationship Uber's driver workforce holds with its rating system. By analyzing the customs and norms within a social group, SCOT practitioners seek to understand the underlying reasons behind a technology's usage and subsequent effects on those social groups (Klein & Kleinman, 2002). In his 1993 paper, "Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology," Social Scientist Langdon Winner, from Rensselaer Polytechnic Institute, criticizes SCOT, describing how the framework often fails to take into account the views of people weren't involved in the creation process of a technology, but must use it regardless. For this paper, SCOT will be used to explore how the circumstances of drivers in different market segments change their view of the driver rating system, and thus have shaped the system itself. Information on how Uber has already updated their system in 2017 will be useful for this (Improved Rating System and Feedback Protection for Drivers, n.d.). Subsequently, various rider segments are analyzed to see how their varying norms or preconceptions of driver segments affect the ratings they give.

This research will be important when considering how to improve rating systems such that they not only align with the interest of the firms or entities which use them, but also with the interests of social groups which are being rated. Understanding how to implement said alignment is important from a business perspective since it will help address principal-agent problems within firms. In Uber's case, this would help ensure its employees deliver the quality service Uber promises without worries about ratings compromising the experience (such as begging for 5 stars). From a social perspective, this research will also help understand how other rating systems can better cater to its target market without ostracizing sub-groups within the market. From a data collection perspective, by understanding how social groups are affected by rating data, analytics can be adjusted to avoid punishing social groups due to contextual factors outside their control (such as low ratings from rider biases).

Research Question and Methods

How have varying social groups influenced the use of Uber's driver rating system and its subsequent effects on drivers?

The two methods used to pursue the research question are wicked problems and network analysis. Wicked problems describe a class of societal issues which are deeply engrained and do not have a clear-cut solution due to the deep complexity of the issue. This method is being used to outline the deeper underlying issues within the social groups (particularly social biases (Hua & Ray, 2018)), as well as the problems Uber wishes to tackle through its driver rating system, namely safety (How Uber Star Ratings Work For Driver-Partners, n.d.). E.K Clemons' 2007 paper on rating system in E-Commerce and Kim, Moravec, & Dennis' 2019 paper on "The Effects of User and Expert Reputation Ratings" will be useful for discussing these underlying issues within social groups. Kostka's 2019 article on Chinese majority acceptance of the Social Credit System will help lean in on how cultural differences in social groups can explain the development of a controversial system that addresses a wicked problem.

Once the circumstances of the social groups are understood, network analysis will be used to explore how the greater social context interacts with Uber and its rating system to produce the disparity in the driver market. This method is useful since it focuses on explaining the relationships between actors within a network, linked together through hierarchies, social group membership, and/or monetary ties. Firmino, Cardoso, & Evangelista's research on "Uber and Surveillance Capitalism by the Global South" (2019) will be helpful in describing the phenomenon of hyperconnected networks closing off certain social groups from Uber's network.

Utz, Matzat, & Snijders' 2009 article on feedback will also help explain how feedback can affect different users in the rating system.

Conclusion

The technical deliverable will be a model for adaptive sensing on phone sensors. This model will allow an application to intelligently turn on and off mobile sensors in order to save battery life while collecting quality data. This deliverable will increase the feasibility of large-scale mobile data collection without impeding the user's day-to-day life. The STS deliverable will be an analytic report on how social groups have influenced Uber's driver rating system and vice versa. Understanding these influences will help Uber develop subsequent actions to take in order to alleviate stress from a portion of its labor market, and possibly create a more stable segment of its market. Ultimately, stabilizing its labor market will allow Uber to align its employees with its vision and thus deliver a higher quality riding service for both drivers and riders.

References

Bowman, C. (2019, October 8). I'm a driver for Uber and Lyft—Here's exactly how much I make in one week on the job. Retrieved October 22, 2019, from Business Insider website: https://www.businessinsider.com/uber-lyft-driver-how-much-money-2019-10

Clemons, E. K. (2007). An Empirical Investigation of Third-Party Seller Rating Systems in E-Commerce: The Case of buySAFE. *Journal of Management Information Systems*, *24*(2), 43–71. https://doi.org/10.2753/MIS0742-1222240203

Cook, J. (2015, February 11). Uber's internal charts show how its driver-rating system actually works. Retrieved October 22, 2019, from Business Insider website: https://www.businessinsider.com/leaked-charts-show-how-ubers-driver-rating-system-works-2015-2

Dickey, M. R. (2014, June 25). Uber's Own Drivers Protested The Company's Policies And Rating System. Retrieved October 3, 2019, from Business Insider website: https://www.businessinsider.com/uber-drivers-protesting-2014-6

Firmino, R. J., Cardoso, B. de V., & Evangelista, R. (2019). Hyperconnectivity and (Im)mobility: Uber and Surveillance Capitalism by the Global South. *Surveillance & Society*, *17*(1/2), 205–212. https://doi.org/10.24908/ss.v17i1/2.12915

How Uber Star Ratings Work For Driver-Partners | Uber. (n.d.). Retrieved October 23, 2019, from https://www.uber.com/drive/resources/how-ratings-work/

Hua, J., & Ray, K. (2018). Beyond the precariat: Race, gender, and labor in the taxi and Uber economy. *Social Identities*, *24*(2), 271–289. https://doi.org/10.1080/13504630.2017.1321721

Improved Rating System and Feedback Protection for Drivers. (n.d.). Retrieved October 3, 2019, from Uber website: https://www.uber.com/us/en/c/180-days/improved-ratings-and-feedback/

Kim, A., Moravec, P. L., & Dennis, A. R. (2019). Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings. *Journal of Management Information Systems*, *36*(3), 931–968. https://doi.org/10.1080/07421222.2019.1628921

Klein, H. K., & Kleinman, D. L. (2002). The Social Construction of Technology: Structural Considerations. *Science, Technology, & Human Values*, *27*(1), 28–52. https://doi.org/10.1177/016224390202700102

Kostka, G. (2019). China's social credit systems and public opinion: Explaining high levels of approval. *New Media & Society*, *21*(7), 1565–1593. https://doi.org/10.1177/1461444819826402

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

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

Philipp, J. (2018, April 30). The Chinese Regime's "Social Credit" Dystopia. *The Epoch Times, New York Ed.; New York (NY)*, p. A1,A14.

Rogers, B. (2015). The Social Costs of Uber. *University of Chicago Law Review Dialogue*, 85–102.

Sparapani, T. (n.d.). Data Dump: Data Collection and Selling is Worse Than You Think. Retrieved October 22, 2019, from Forbes website: https://www.forbes.com/sites/timsparapani/2019/10/22/data-dump-data-collection-and-selling-is-worse-than-you-think/

Summers, L. H. (n.d.). Larry Summers: Data collection is the ultimate public good. Retrieved October 22, 2019, from Washington Post website: https://www.washingtonpost.com/news/wonk/wp/2016/04/04/larry-summers-data-collection-is-the-ultimate-public-good/

Utz, S., Matzat, U., & Snijders, C. (2009). On-line Reputation Systems: The Effects of Feedback Comments and Reactions on Building and Rebuilding Trust in On-line Auctions. *International Journal of Electronic Commerce*, *13*(3), 95–118. https://doi.org/10.2753/JEC1086-4415130304

Winner, L. (1993). Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology. *Science, Technology, & Human Values*, *18*(3), 362–378. https://doi.org/10.1177/016224399301800306