Thesis Portfolio

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

(Technical Report)

Social Challenges of Uber's Driver Rating System

(STS Research Paper)

An Undergraduate Thesis

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

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

Ian Tucker
Spring, 2020

Department of Systems and Information Engineering

Table of Contents

Sociotechnical Synthesis

Today, companies are harnessing the power of user data in order to drive success in their business, particularly for targeted advertising and innovation in business models. Day by day, modern companies not only utilize user data analytics to curate results for their users, but also to learn more about different user groups and the decisions they make in regards to these companies. Both the technical and STS projects outlined in this portfolio explore different aspects of data and data analytics, and how they currently affect or will affect the general populous. The technical research explores the relationship between a single user's data and their health, and how modeling can be used in order to provide information on a user's wellbeing without encroaching on their wellbeing in real-time. The STS portion of research explores the relationship between data created not by a single user, but by two users, a driver and a rider for Uber, in the form of a star rating. The resultant power relationship between the two users caused by this data and the resultant long-term network effects are discussed. In effect, the results of these research provide key insights on micro and macro-scale effects of data on users and their human relationships.

The purpose of the technical project is to provide a framework for how smartphone sensor data can be collected, cleaned, stored, and modeled to effectively predict human states. Outlined in the paper is a three-week long study involving an application named Sensus, which collects and aggregates user smartphone data. Differing models for how the application collects data were tested in order to determine which strategy collects data of the highest quality with minimal battery consumption. Sensing strategies range from infrequent sampling to continuous sampling. A final, dynamic data collection strategy was then explored, using a machine learning model trained on existing data collected from 220 participants to determine the best times to

trigger sensor sampling. Results of this study will include 1) extraction of model features that deliver maximized data quality with minimized battery consumption as compared to pre-existing baseline models, 2) implementation of context-driven modeling of user smartphone data on user's contextual environment, and 3) customization of a time-series database for optimized data queries used in metadata visualizations. The adaptive sensing models produced could be used in future large population studies that efficiently examine patterns of behavior in multiple individuals over extended periods to identify disease indicators present in an average user's daily life.

The purpose of the STS research is to explore the social ramifications of Uber's star rating system on Uber's workforce-user base and the schism it creates in its labor-market segment. This research specifically seeks to answer how varying social groups have influenced the use of Uber's driver rating system and its subsequent effects on drivers. Methods of Network Analysis and Wicked Problem Analysis are used to reconstruct the complexity of the issues faced between social groups, namely that of social bias, and subsequently explore their connections to the rating system and the larger Uber environment. The Social Construction of Technology framework is used to further explore how these engrained societal viewpoints of varying driver segments affect the use of the rating system. Findings from this research suggest that the racial biases of riders and the lean application structure of Uber plays into creating an asymmetrical information space that leads passengers to rate their riding experience based on comparison to (uncommunicated) expectations as opposed to the objective quality of provided service itself. The resultant network structure centralizes power in part-time non-immigrant drivers, which, inadvertently, allows them easier access to increased ratings over the long run. This research is important in learning how to generate social rating systems on larger scales that

avoid unnecessary subjectivity or account for it such that specific user groups are not disenfranchised by future systems.

These two projects were separated in order to understand different perspectives of how data can affect a user's daily life. In one, I learned about how data can be used to gain insights on a user on an individual level. In the other, I learned how data affects multiple users and their relationships. By working on them in parallel, I gained substantial insight in how human psychology affects data collection on both the micro and macro-scale and the interconnections between those effects. Learning about how asymmetric information and expectations affect riders' ratings in Uber was translated to the technical project in how to keep users informed and ensure they answer questions as honestly and objectively as possible. Exploration of models for human behavior supported research on the construction of a social network model for Uber ratings. Ultimately, this project has shaped my view on the interconnected nature of data and human society and how to best navigate this growing field for the seamless integration of the two.