Data Integration for Enhanced Movement Behavior Analysis

Sri Surya Krishna Rama Taraka Naren Durbha, Hedyeh Mobahi, Janusz Wojtusiak
George Mason University, Fairfax, VA, USA

Introduction

Understanding how people move and behave has become a vital tactic in the fight against the transmission of infectious diseases, especially during worldwide pandemics. There are many lessons learned from the COVID-19 pandemic, during which movement restrictions and policies have been implemented variably across regions. Many of these works track large volumes of people using aggregated datasets[1]. Unfortunately, the aggregated nature of the datasets does not allow them to be linked with other datasets, limiting the applicability of analyses. This presentation reports on integration of four datasets collected as part of a project aimed at understanding movements of people during different phases of COVID-19 pandemic.

Methods

During the COVID-19 pandemic a study was designed to collect and integrate GPS, WiFi, health and vaccination status as well as a large survey of attitudes towards social distancing of enrolled participants from George Mason University. Study was rolled out in two stages as the design evolved. The first stage included collecting health status and geolocation from participants who signed up for the study. Participants in cohort 1 consented to share their GPS data either through an iPhone or Android applications. In addition, the university required all individuals participating in any on-campus activities and all residential students to complete daily symptom and COVID-19 diagnosis questionnaire, called Mason HealthCheck. The tool was implemented in August 2020, building on experimental tools that had been developed and used since March 2020. As the pandemic evolved, the study enrolled cohort 2 in which participants consented to sharing also WiFi data. The study collaborated with the university's Information Technology Services (ITS) team to retrieve WiFi connection log data for these set of participants. The WiFi data consisted of participant ID and time/location of the connection on the university's Fairfax campus. As COVID-19 vaccination progressed, a survey was distributed to the study participants to gather information on their vaccination status and capture behavioral data. All the above datasets included participant IDs, which allowed for integrating data into one database. All identifiable participant IDs were replaced with random unique identifiers before proceeding to further analysis.

Results

A total of 376 participants signed up for the study. However, due to non-compliance, technical issues and/or discontinuation, GPS data is available for only 196 individuals covering the timeframe from 05/06/2020 to 02/26/2022. Further, 326 participants actively reported their health status, 132 participants reported their vaccination status and WiFi data was available for 93 participants. The numbers of participants in each dataset can be seen in Figure 1, notably complete integrated data is available for 61 study participants.
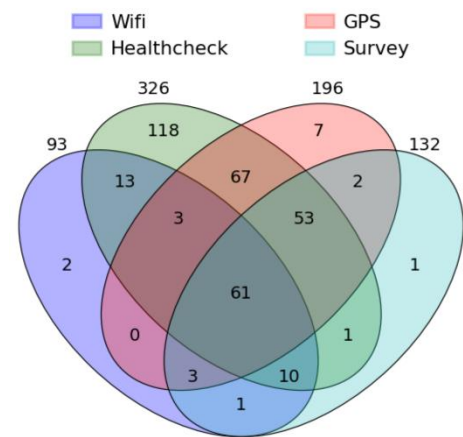

Figure 1: Numbers of participants with data in four datasets.

Discussion and Conclusion

With this integrated data, we can further investigate movement patterns, contact trace if we detect any ill health reporting and inform individuals who might be at risk to take additional precautions with fewer resources. Of particular interest is the trend in specific locations that people visit outside of home (i.e., store, restaurant) in the context of being symptomatic or asymptomatic (HealthCheck), as well as their varying attitudes towards COVID-19 and social distancing (Survey). To summarize, the integration of these diverse data sources enhances the accuracy, timeliness, and effectiveness of contact tracing, thereby empowering public health authorities to mitigate the spread of infectious diseases more efficiently.

References

1. Owuor, I. and Hochmair, H. H.: Use of SafeGraph visitation patterns for the identification of essential services during COVID-19, AGILE GIScience Ser., 4, 36, https://doi.org/10.5194/agile-giss-4-36-2023, 2023.
2. Wojtusiak J, Bagchi P, Durbha SS, Mobahi H, Mogharab Nia R, Roess A. COVID-19 symptom monitoring and social distancing in a university population. Journal of healthcare informatics research. 2021 Mar;5:114-31.