Agent-Based Simulation of Human Mobility Using High-Resolution Foot-Traffic Data

The 2021 Aspiring Scientists' Summer Internship Program

DARIUS KIANERSI, FIONA ZHENG, Justin Elarde, Fahad Aloraini, Amira Roess[3], Hamdi Kavak[1], Tim Leslie[2], Andreas Züfle[2], Taylor Anderson[2]

[1]Department of Computational and Data Sciences, [2]Department of Geography and Geoinformation Science, [3]College of Health and Human Services

Introduction and Purpose

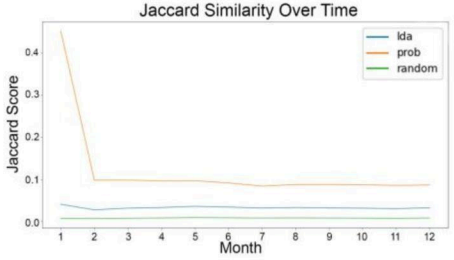
Accurate models for human mobility can be used to simulate autonomous transportation, urban planning, and disease transmission, particularly within the scope of the COVID-19 pandemic. With realistic models for disease transmission, interventional and preventative policies can be implemented to affect human mobility and subsequently mitigate disease spread. Although these models of human mobility offer several advantages, their implementation is limited by contemporary models which lack the sufficient data to make educated inferences. As a result, human mobility is largely random in these simulations and lacks consideration of geographic constraints and visitor demographics.

Methods

We acquire data from SafeGraph, which provides smartphone location data at the Census Block Group (CBG) level and records user visits to various Points of Interest (POIs). We use an Agent-Based Model (ABM) to model mobility, in which agents build schedules daily and interact with other agents before returning home. We compare the accuracies of three statistical models to infer human mobility: a probabilistic model and Latent Dirichlet Allocation (LDA) model, which are both data-driven, and a random model, in which agents may visit any POI notwithstanding visitor demographics or other constraints.

Results

Model	Jaccard Similarity Score (yearly average to ground-truth)
Probability	0.123
LDA-649	0.092
LDA-75	0.048
LDA	0.033
Random	0.008



Conclusions and Future Work

- The probability model performed better than the LDA and random models
- The random model performed significantly worse than the data-driven models
- All data-driven models performed better on the month they were trained on (January), and their accuracies stabilized over time
- LDA model began to approach the probability model as topics increased
- Probability model may overfit to the population of SafeGraph visits
- Future work is to train the models on a reduced population of POI visits

Major Citations

Pesavento, J., Chen, A., Yu, R., Kim, J. S., Kavak, H., Anderson, T., & Züfle, A. (2020, November). Data-driven mobility models for COVID-19 simulation. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities (pp. 29-38).*

SafeGraph. 2020. Patterns. Accessed August 9, 2021. https://docs.safegraph.com/docs/monthly-patterns

Acknowledgements

This work is supported by National Science Foundation Grant #2109647 titled "Data-Driven Modeling to Improve Understanding of Human Behavior, Mobility, and Disease Spread". This work is supported by the George Mason University Aspiring Scientists Summer Internship Program (ASSIP).