Regularities in Mobility Patterns at Different Time Scales

Evidence from "Digital Trace Data" in the U.S. and Senegal

Lee Fiorio – Department of Geography
University of Washington
fiorio@uw.edu

CENTER FOR STUDIES IN DEMOGRAPHY & ECOLOGY

Abstract

Recent decades have seen an explosion in the quantity of behavioral data generated by human interaction with digital devices. A growing body of literature has focused on the value and potential pitfalls of leveraging these "digital trace data" to analyze social processes including human migration and mobility, but blind spots remain. One challenge, well known to migration scholars, is to standardize and compare different kinds of movements across different time and geographic scales. In this paper, we develop a methodology for parsing the population-level migration signal from individual-level point-in-time data using flexible time-scales. We propose a stochastic model for simulating patterns in digital trace data and test it against three datasets: geo-tagged Tweets and Gowalla check-ins in the U.S.; cell phone call detail records in Senegal. Similar patterns observed across all three empirical datasets demonstrate the utility of our approach for studying migration via digital trace data.

1. Motivation

Individual-level location-in-time data are increasingly common. They are contained in:

- cell phone call detail records
- social media posts
- metadata associated with smart phone applications
- log-ons to emails
- other sources

These so-called "digital trace data" come in large collections and are simple in form:

<unique individual ID, timestamp, location>

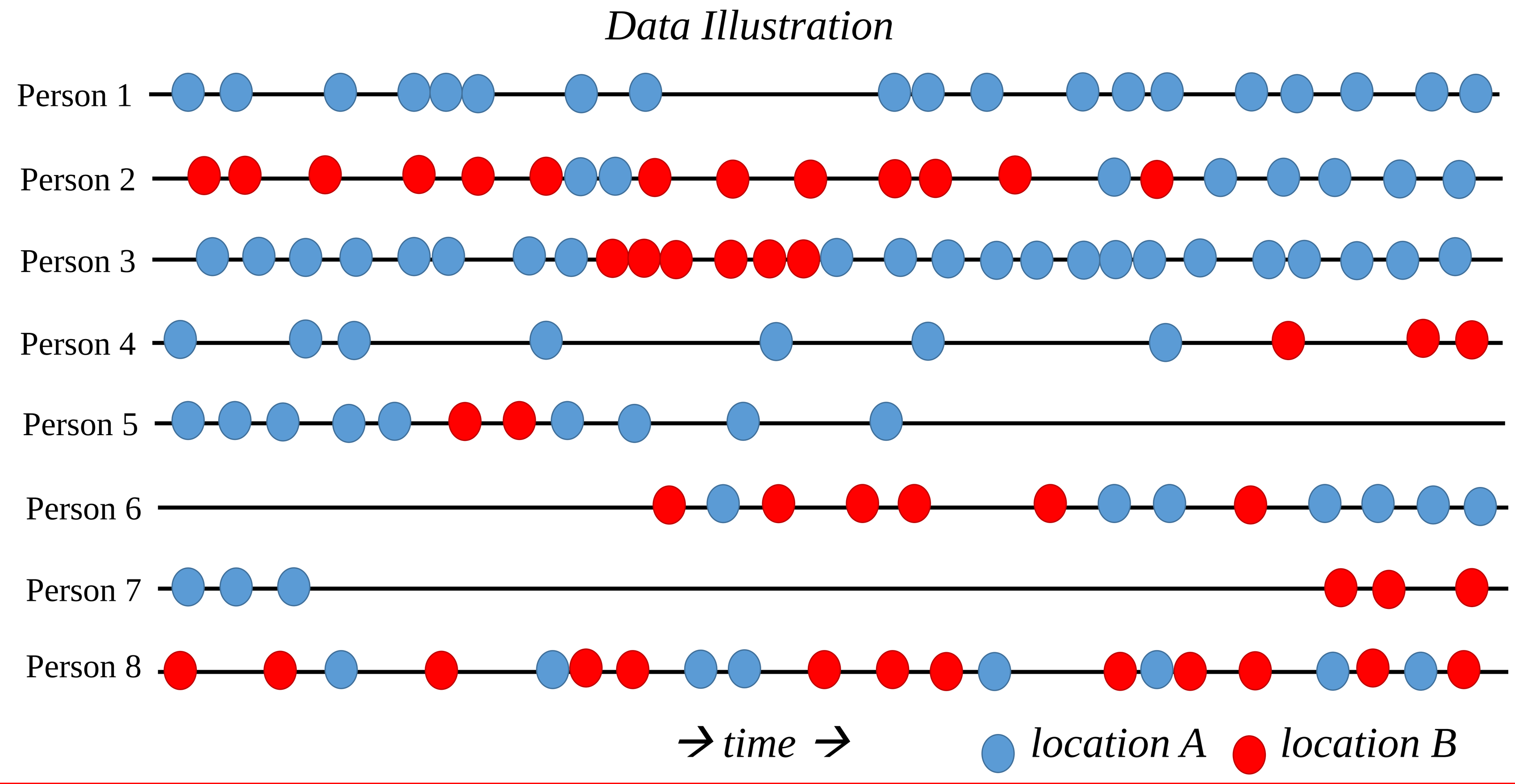
But there are no standards for measuring migration signal in these kinds of data.

2. Objective and Strategy

To develop a methodology for measuring bilateral migration flows from "digital trace data," we introduce three simple concepts, adapted from migration literature, that can be implemented flexibly to identify regularities in mobility patterns at coarse geographical scales:

1. ***Buffer*** – period for inferring place of residence
2. ***Interval*** – period between two inferred places of residence (different place of residence at each end of the interval indicates migration transition)
3. ***Start*** – the specific time associated with the earlier place of residence

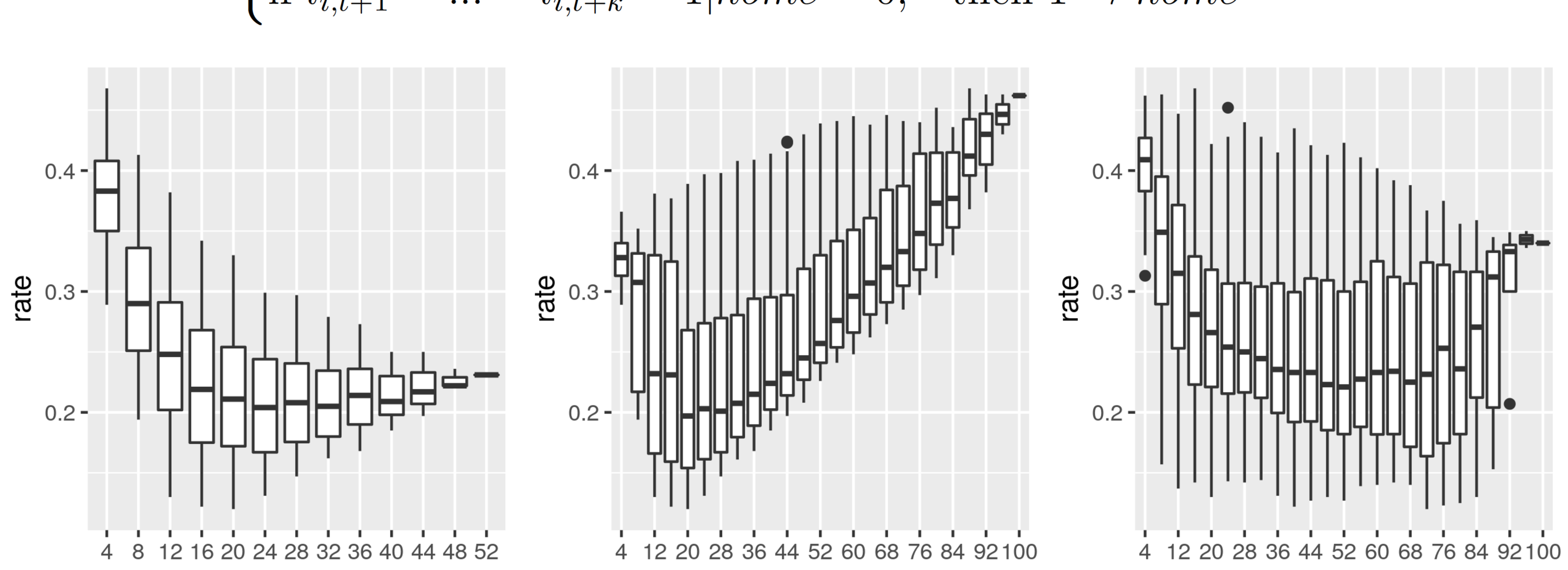
Data Illustration



3. Simulation Model

$$\{l_{i,1}, l_{i,2}, \ldots, l_{i,t}\} \qquad P(l_{i,t}|home=1) = \begin{cases} p, & \text{for } l_{i,t}=1 \\ 1-p, & \text{for } l_{i,t}=0 \end{cases}$$

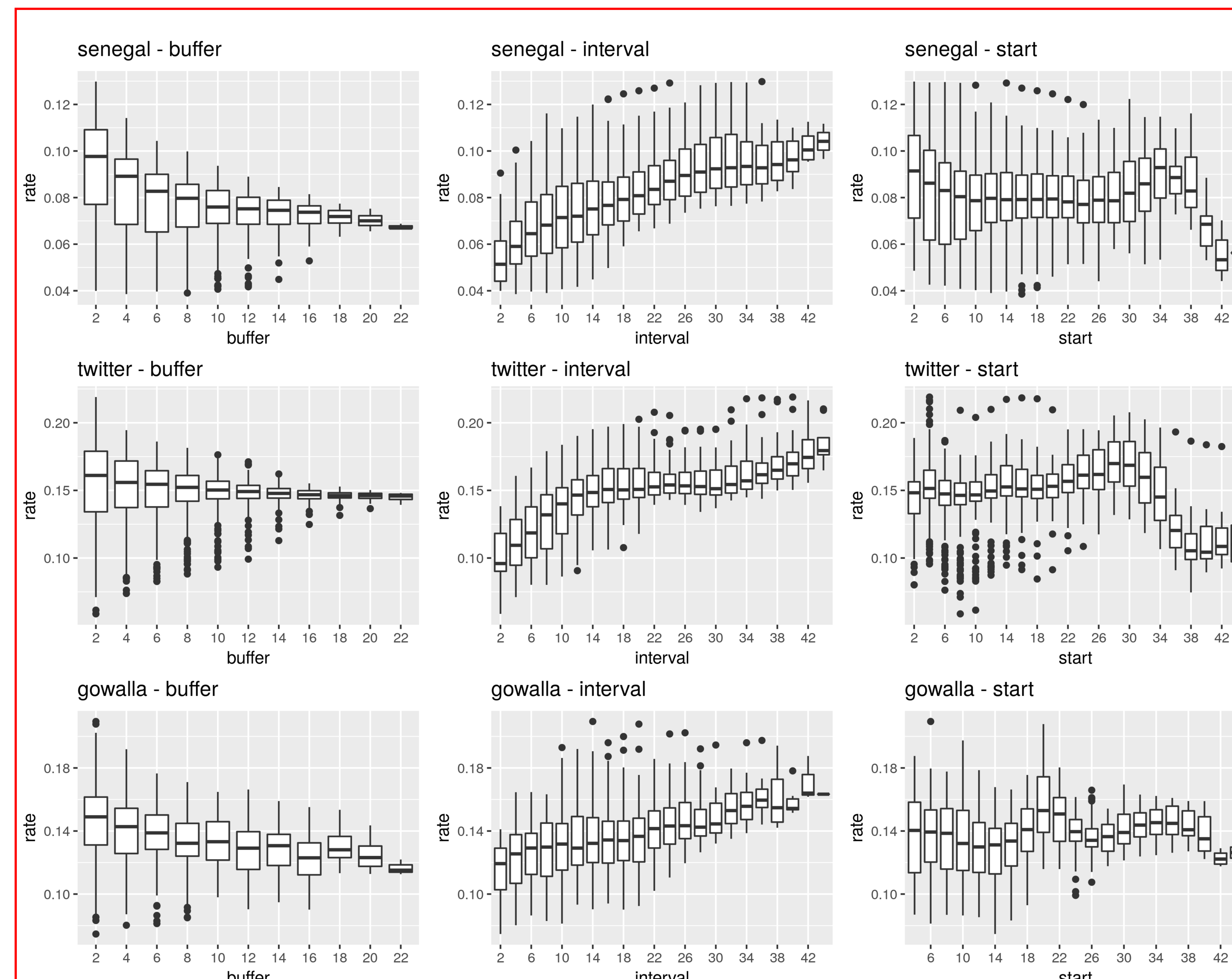
$$\begin{cases} \text{if } l_{i,t+1} = \ldots = l_{i,t+k} = 0|home=1, & \text{then } 0 \rightarrow home \\ \text{if } l_{i,t+1} = \ldots = l_{i,t+k} = 1|home=0, & \text{then } 1 \rightarrow home \end{cases}$$



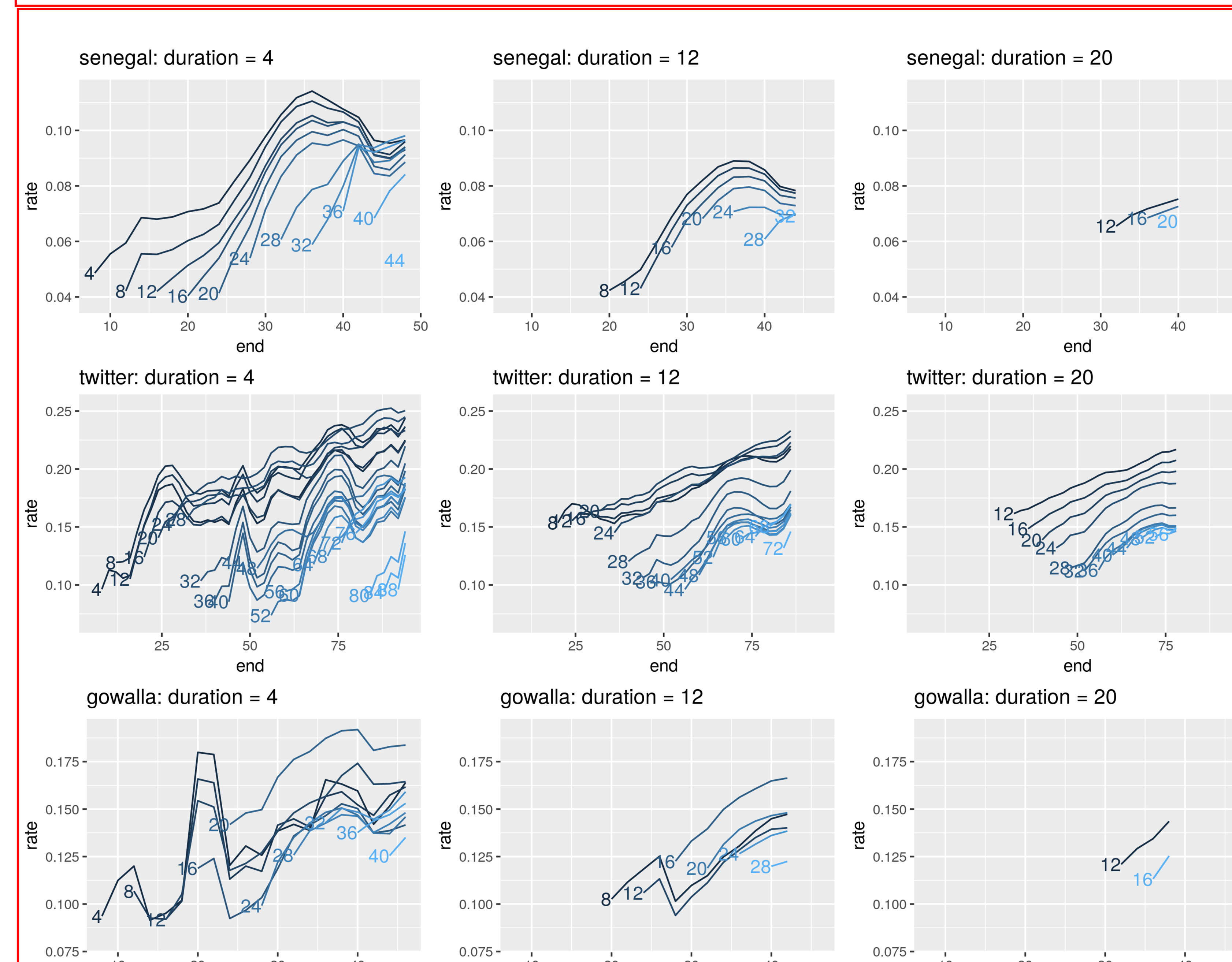
4. Data

- **Senegal *Orange* Cellphone Call Detail Records:** 561 million records from 140k users; 2013 calendar year; 14 administrative regions
- **U.S. Geo-tagged Tweets:** 570 million records from 2.9 million users; 2011 to 2016; 9 U.S. Cesus Divisions
- **U.S. Gowalla check-ins:** 6.4 million records from 109k users; 2009 to 2010; 9 U.S. Census Divisions

5. Empirical Findings



Distribution of migration rates with respect to buffer, interval and start in Senegal-Orange, U.S.-Twitter, and U.S.-Gowalla datasets



Migration rates plotted by start, holding duration fixed

Each line is a set of rates estimated from a particular start, extending the interval forward in time

6. Conclusion

This analysis demonstrates regularities or an internal logic to the migration signal within "digital trace data." Similar to the patterns we modeled in our simulation, migration rates are increasing with respect to interval. There is a negative relationship between duration and the variance in the rates observed. Though they were generated in different contexts (U.S.A., Senegal) and with different underlying behavior (cell phone calls, geo-tagged tweets, location check-ins), the three datasets explored here exhibit a regularity with respect to the concepts – buffer, interval, start – employed here. Further work should be conducted to theorize the relationship between mobility and migration.