Understanding Mobility around NUS using Internal Shuttle Bus (ISB) WiFi Data

A GE3238 project by:
James Tan (A0093894L)
Teo Mingjie (A0154254J)
Eunice Ng (A0157195W)
Eva Aldridge (A0156616B)
Desmond Soh (A0154215N)

Abstract

WiFi connection data has the potential to reveal travel trajectories. Given the availability of such data based on NUS ISB WiFi connections, crowd trajectory data can be mined for information on movement. We propose a versatile workflow to automate data processing towards the identification of proximity to point-of-interests (POIs). The end product of the workflow can be visualised to suggest spatial and temporal crowding patterns on NUS ISBs.

1. Introduction

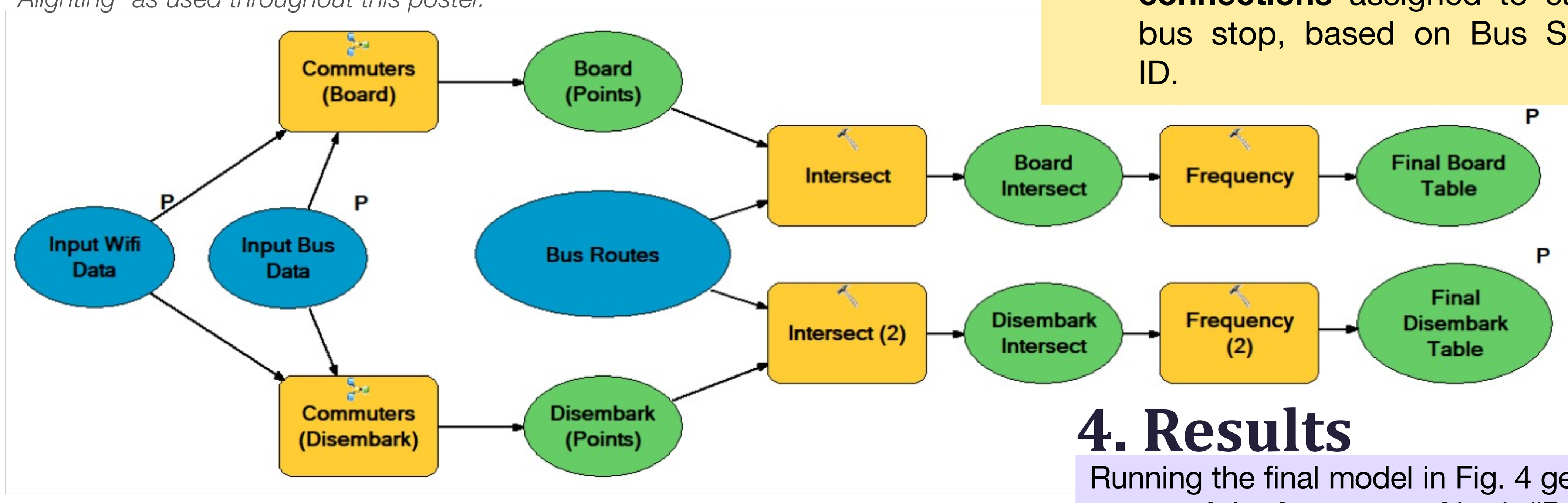
Since 2013, NUS has tried to manage crowd flow on its Internal Shuttle Bus (ISB) services (Sim, 2013). This requires crowd size data which has traditionally been extrapolated from NUS Office of Campus Amenities (OCA) ridership surveys. With the main aim of understanding the mobility of ISB users on a typical day, the recent deployment of wireless mesh vehicular technology potentially provides better data for measuring crowd sizes. This is because it utilises the connections of bus riders' devices (i.e. smartphones) to the bus WiFi routers.

However, the datasets in its original state requires prior processing before it can provide useful insights. Thus, through **ArcMap's ModelBuilder** function, individual WiFi and ISB datasets were cleaned and processed to allow easy visualisation. This presents a versatile and editable method purposed towards the optimisation of the ISBs and reducing overcrowding. We present a demonstration of the workflow using the case study of **A1 Bus WiFi connections in AY2017/18 Semester 2 Week 1 Day 1**.

2. Literature Review

- Smartphone-based travel survey supports data collection initiatives for transport modelling purposes, since smartphones are **increasingly ubiquitous** and are **versatile loggers** (Cottrill et al., 2013)
- WiFi activity is effective in providing location data as it is a **scalable solution** which enables **unrestricted tracking of many people at a low cost** (Chilipirea et al., 2018). Thus, WiFi routers are **high-resolution outdoor positioning tools** (Sapiezynski et al., 2015) for measuring crowd size.
- WiFi data has **low positional accuracy relative to GPS data** (Chilipirea et al., 2018). Merging WiFi-based data with GPS data thus requires a compromise in positional accuracy.

Figure 4: Complete Workflow Model. "Disembark" in this model refers to "Alighting" as used throughout this poster.



3.4 Final Workflow Model

Final products of the model are point-feature classes of individual board and disembark data linked to unique time information, as well as an aggregated count tied to each bus stop along the A1 route. All of this is summarised into the **complete workflow model** (Fig. 4).

5. Assumptions & Limitations

- Assumes any connections in the time between two bus stops belongs to the first bus stop.
  - Accounts for time lag in boarding bus and connecting to WiFi.
  - Neglects possibility of non-passengers connecting via proximity (i.e. traffic jam).
- Using 4 d.p. positional accuracy cannot account for WiFi connections greater than 11.1m from bus route.
- Representativeness of data is limited:
  - Cannot account for commuters with no login-access to NUS WiFi network
  - Cannot account for NUS Staff or Student commuters who do not login to WiFi onboard the bus.

3. Methodology & Datasets

3.1 Data Acquisition

WiFi and ISB datasets were obtained from **NUS data commons.** To supplement data analyses, A1 bus stop coordinates were digitised using **Google Maps visual checking**.

Info needed for respective attribute tables:

WiFi	Node ID; Start & End for GPS Time & Coordinates
Bus	Node ID; Vehicle No.; GPS Time; Latitude & Longitude

3.1 Data Cleaning

- Model created as seen in Fig. 1.
- Rounding coordinates data to 4 d.p.
- Creating Unique ID based on coordinates and time data.
- Conversion of date/time field from UTC format to compatible format for Time Slider visualisation*

**Time Slider out of project's scope, may be considered for future works*

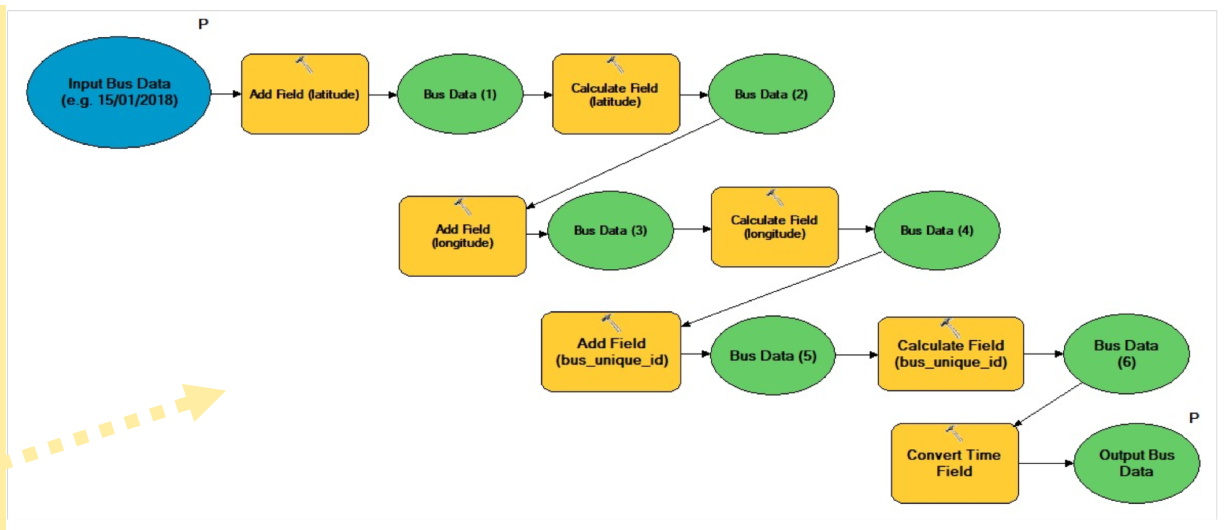


Figure 1: ISB data cleaning model created

3.3 Data Processing

Data processing in ModelBuilder was subdivided into 3 steps:

I. Identification of individual bus routes:

- A1 bus stops and ISB: latitude and longitude data were **concatenated**.
- New concatenated fields for both bus stop and ISB data tables were then **joined,** resulting in **A1 bus routes.**
- A1 routes further narrowed down based on unique bus stops.

II. Separate "Boarding" and "Alighting" data:

- **2 models** were created for "boarding" (Fig. 2) and "alighting" bus commuters respectively.
- ISB and WiFi data: latitude, longitude, and time were **concatenated** for **uniqueID.**
- ISB and WiFi joined on matching records.
- **Table Select** using **A1 NodeID** (e.g. represented as '2074' in this project) keys to obtain ONLY A1 information.
- **Make XY Event Layer** of both "Boarding" and "Alighting" for separate point-feature classes.

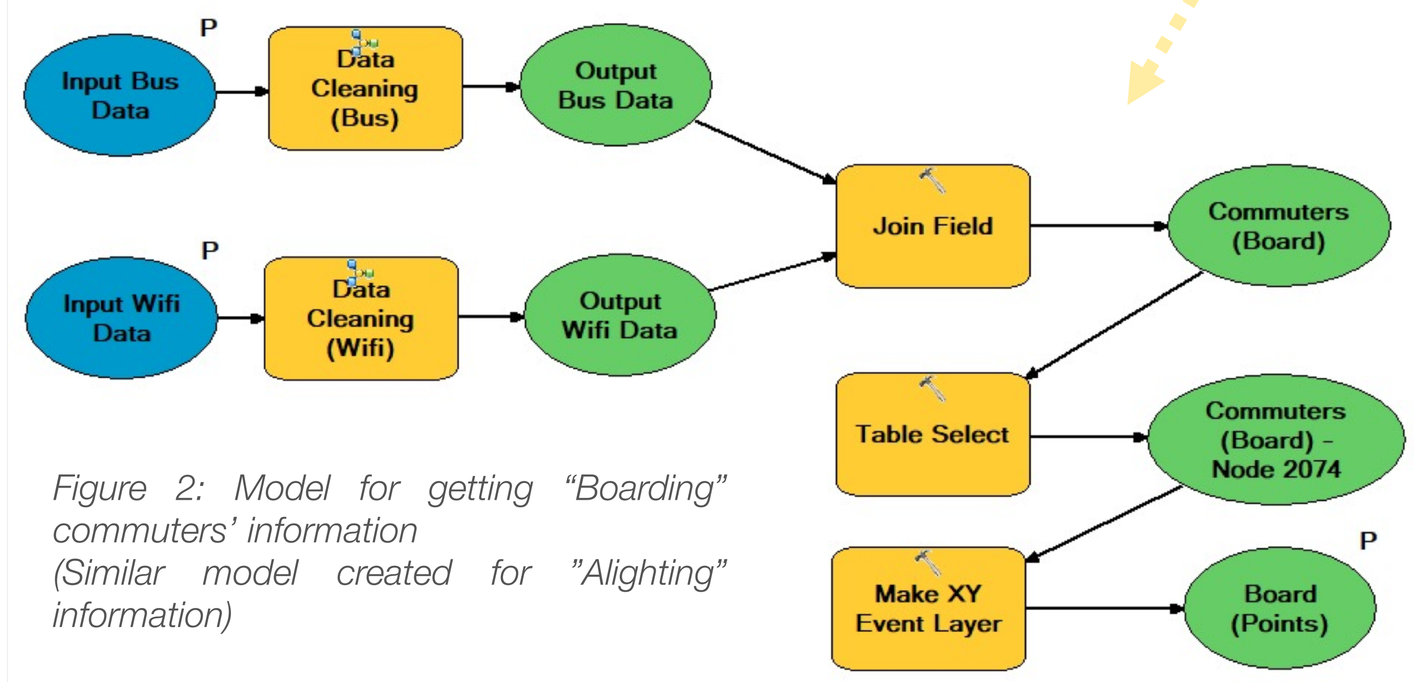


Figure 2: Model for getting "Boarding" commuters' information (Similar model created for "Alighting" information)

III. Spatial assignment of "Boarding" and "Alighting" data points:

- Using **A1 bus stop** locations, entire A1 bus route was digitised (Fig. 3):
  - Assign polygon between Bus Stop 1 (i.e. YIH) and Bus Stop 2 (i.e. CLB) as Bus Stop 1 'catchment area' for lagging WiFi connections.
  - **Polygon retains Bus Stop ID.**
- **Intersect** digitised route with "Boarding" and "Alighting data separately, with tolerance of 5m.
- Count the **frequency of WiFi connections** assigned to each bus stop, based on Bus Stop ID.

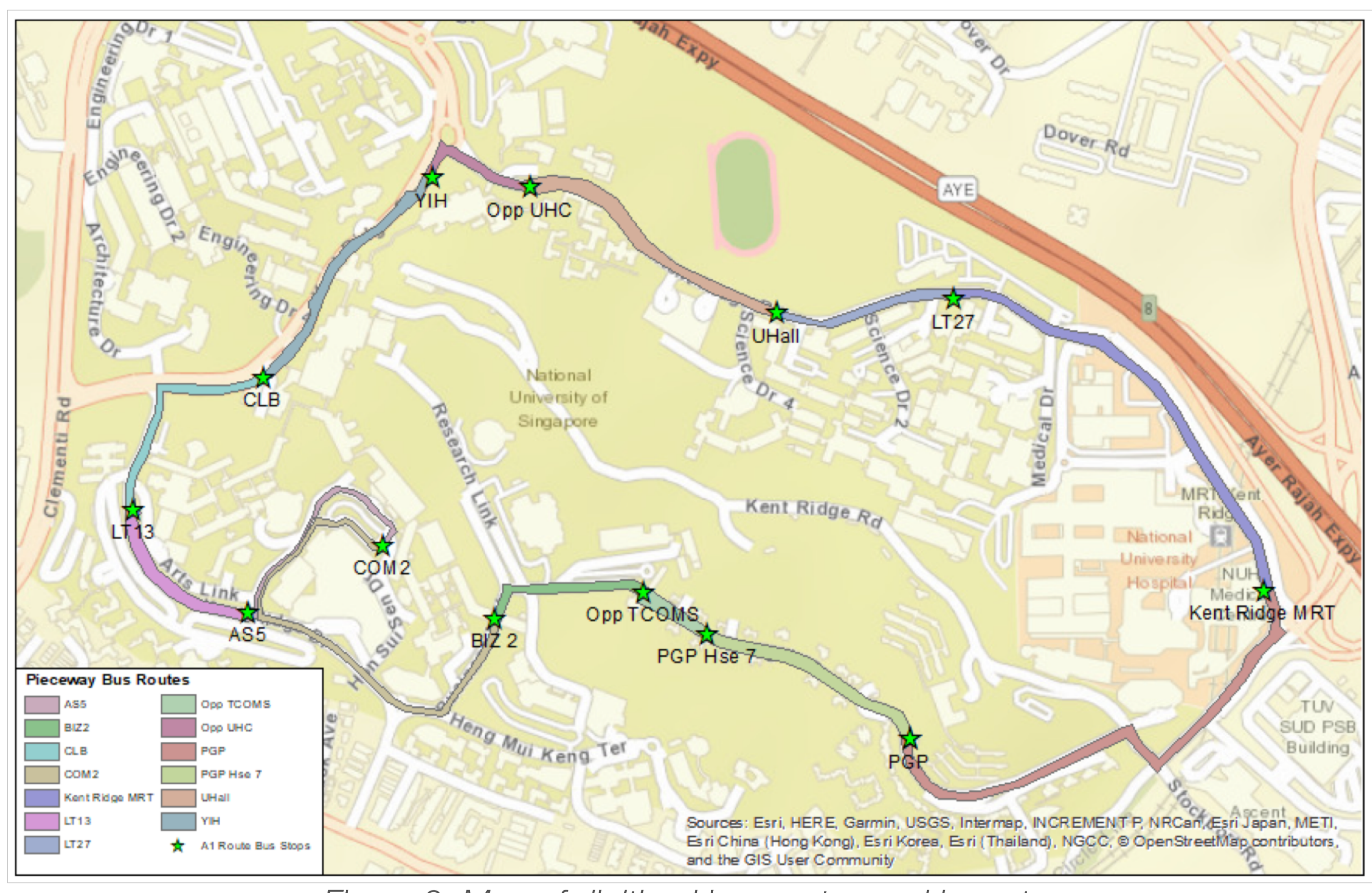


Figure 3: Map of digitised bus routes and bus stops

4. Results

Running the final model in Fig. 4 generated a map of the frequency of both "Boarding" and "Alighting" commuters taking A1 on 15/01/2018 (Fig. 5). The results tally with existing knowledge of popular embarking (KR MRT) and disembarking (PGP terminal) sites.

Future work for this project can include users applying model to visualise mobility patterns and conduct spatial analysis on a larger scale from resulting model outputs. Future iterations may also include unique NUS web ID to consider passengers with >1 connectable device, as well as location and accelerometer data to complement bus stop site information to detect transitions for better map display. Lastly, future projects could carry on from this to use Time Slider and multiple routes to provide a comprehensive overview of crowd movement throughout the day.

Frequency of People Boarding and Alighting at NUS Bus Stops on Route A1 (15/01/2018)

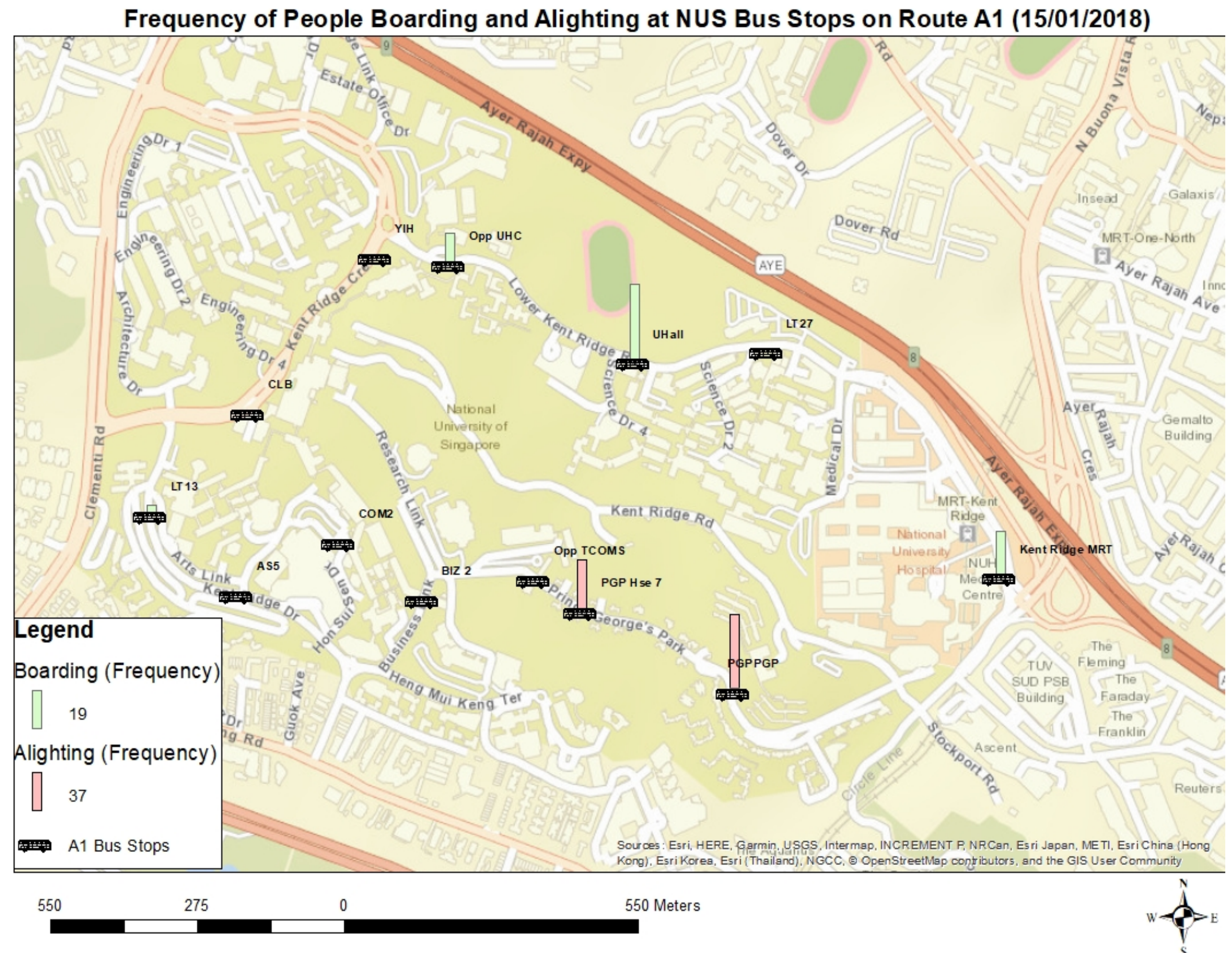


Figure 5: Map displaying results from final model

References:

- Chilipirea, C. et al. (2018). Identifying Stops and Moves in WiFi Tracking Data. *Sensors, 18*(11), 4039. doi:10.3390/s18114039.
- Cottrill, C. D. et al. (2013). Future mobility survey: experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board, 2354*(1), 59–67. https://doi.org/10.3141/2354-07.
- Sapiezynski, P. et al. (2015). Tracking human mobility using wifi signals. *PLOS ONE, 10*(7), e0130824. https://doi.org/10.1371/journal.pone.0130824.
- Sim, R. (2013, September 30). NUS looking for solution to shuttle bus squeeze [Text]. Retrieved from The Straits Times website: https://www.straitstimes.com/singapore/nus-looking-for-solution-to-shuttle-bus-squeeze-0