Transportation Analytics and Optimization with 802.11 Probe Tracking

Syedur Rahman, Jason Scatena, Daniel Su, Peter Zhang
Supervised by Professor Naryan Mandayam

Motivations and Objectives

• The Rutgers University bus system runs into issues of scale and inefficiencies with its massive rider volume.
  • See “Transit Inefficiencies” to the right
• With the majority of riders owning smartphones, we propose using WiFi tracking to obtain rider movement data and transit volume in high traffic zones.
• We want to use predictive analytics and other methods to extract

Methodology

We used Debian based Raspberry Pi’s with USB WiFi adapters to act as data collection stations.

Our tracking methodology is based on the 802.11 probe request. This packet is constantly being broadcasted by modern smartphones at regular intervals in order to identify nearby WiFi access points. We set our network adapter into monitor mode in order to capture all probe requests emitted in a particular geographic zone.

The captured probe requests are then parsed to acquire a timestamp, signal strength, and unique device id.

Finally we associate each collection with its geographic location and collate our data.

To the right is an image of a R-pi data collection station. We place one of these in each of our chosen high traffic locations.

Below the R-pi is a battery pack that we use when deploying to remote locations without a power outlet.

Challenges

Our project originally was much bolder and we wanted to obtain a full representation of Rutgers student movement. We wanted to use the existing Rutgers University wireless infrastructure. However we ran into a large amount of red tape in trying to coordinate this.

We also faced many challenges in deploying our Raspberry Pi stations. They needed to be placed at a location where they would get an accurate reading of the designated zone. Additionally we had to consider power, shelter, and student tampering. We also wanted to avoid situations where these devices would be identified as security threats.

Another issue was dealing with a massive dataset and scalability issues. We gathered around 10,000 – 30,000 requests per day from 2,000 – 5,000 unique devices.

Acknowledgements

Prof. Naryan Mandayam and WINLAB for their support

Transit Inefficiencies

• We developed a mechanism to determine optimal bus allocations as a percentage of total buses
  • Our algorithm calculates a marginal benefit for each additional bus added to a route
  • Ideal allocation percentages are created by “adding” a set number of busses at each time slice to the route with the most to gain
• To the right are our results given a 3 campus system the stops are:
  • Scott Hall (SH)
  • Livingston Student Center (LSC)
  • Allison Road Classrooms (ARC)
• In our results the A bus (Busch Campus) is being underrepresented due to the limited traffic flow from the ARC bus stop.

Results

Campus Traffic Flow

16:15

+118 165 -143

Predictive Analysis Results

Ideal Allocation of Buses by Route

Route

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

0 1 2 3 4 5 6 7 8 9 10 11 12

Time of Day as % of Slices

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

0 1 2 3 4 5 6 7 8 9 10 11 12

Time of Day as % of Slices