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

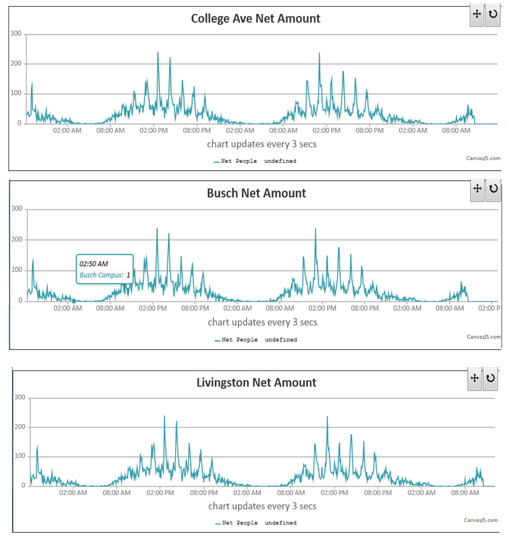
Prof. Narayan Mandayam and WINLAB for their support

Results

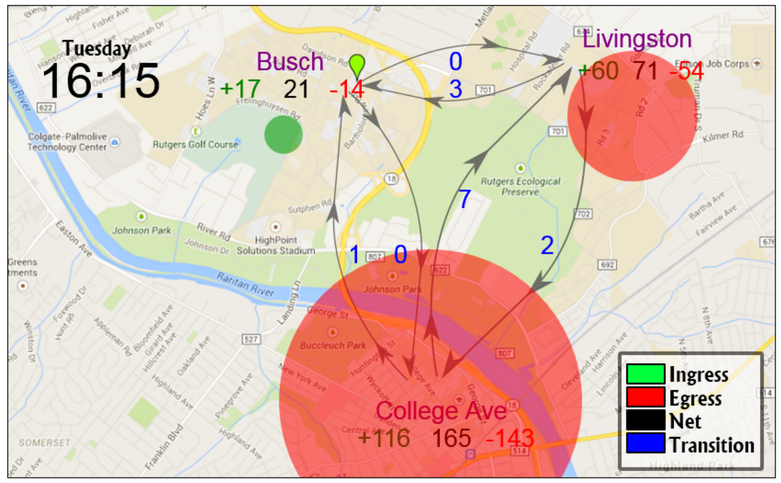
Parsed Data

Time dB MAC

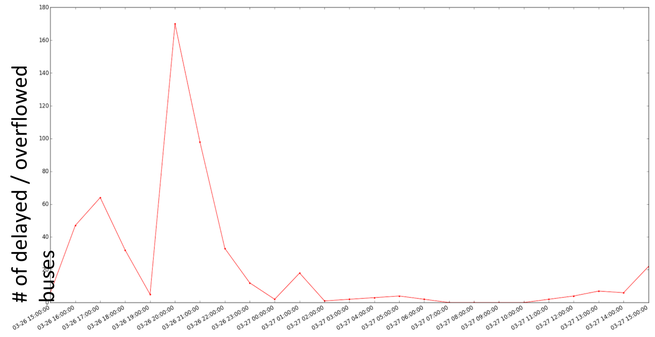
```
1427381021 -87dB 0a:11:06:
1427381021 -87dB 0a:11:06:
1427381021 -93dB 4c:7c:5f:
1427381021 -35dB c0:bd:d1:
1427381021 -35dB c0:bd:d1:
1427381022 -91dB dc:9b:9c:
1427381022 -89dB dc:9b:9c:
1427381024 -85dB a0:5b:78:
1427381024 -85dB a0:5b:78:
1427381024 -85dB 66:de:12:
1427381024 -85dB 66:de:12:
1427381024 -87dB a0:5b:78:
1427381025 -85dB a0:5b:78:
1427381025 -77dB a0:5b:78:
1427381025 -81dB a0:5b:78:
1427381025 -95dB 40:0e:85:
1427381025 -97dB 40:0e:85:
1427381026 -91dB 88:63:df:
1427381027 -81dB 6c:3e:6d:
1427381027 -81dB 6c:3e:6d:
1427381027 -87dB 9c:d9:17:
1427381027 -85dB 9c:d9:17:
1427381027 -87dB 9c:d9:17:
1427381027 -87dB 9c:d9:17:
1427381027 -79dB 72:8c:53:
1427381027 -79dB 72:8c:53:
1427381027 -87dB 9c:d9:17:
1427381027 -89dB 9c:d9:17:
1427381028 -89dB 9c:d9:17:
1427381028 -87dB 9c:d9:17:
1427381028 -89dB 3c:d0:f8:
1427381028 -95dB 9c:d9:17:
1427381028 -97dB 9c:d9:17:
1427381028 -95dB 9c:d9:17:
1427381028 -93dB 9c:d9:17:
1427381028 -95dB 9c:d9:17:
1427381030 -87dB 26:a7:6f:
1427381030 -87dB 26:a7:6f:
1427381030 -35dB c0:bd:d1:
1427381031 -91dB a0:5b:78:
1427381031 -89dB a0:5b:78:
1427381031 -89dB f0:24:75:
1427381031 -75dB a0:5b:78:
1427381031 -75dB a0:5b:78:
1427381031 -87dB f0:24:75:
1427381031 -73dB a0:5b:78:
1427381031 -75dB a0:5b:78:
1427381031 -91dB dc:9b:9c:
1427381032 -91dB 1c:1a:c0:
```



Campus Traffic Flow



Transit Inefficiencies



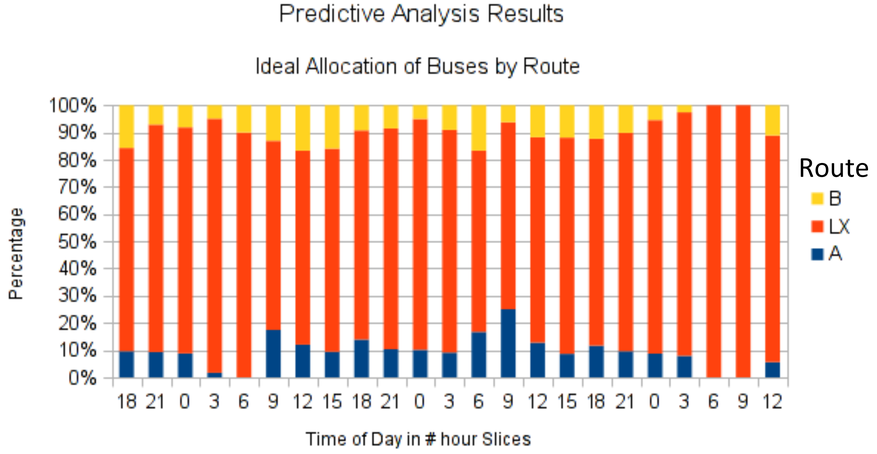
Time of Day

- "Transit Inefficiencies" shows the volume of riders who experience unusually long waiting times due to either delays or overflowed buses.
- There is a cyclic trend on the College Ave campus.
- The supply of buses that were on time consistently did not meet the demand of riders on weekdays at
  - 3:00 – 6:00 PM
  - 7:00 – 9:00 PM.

- We can compensate this deficiency by reallocating buses to different routes

- 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.

Predictive Analysis Results



RUTGERS