Exploring Pedestrian Bluetooth and WiFi Detection at Public Transportation Terminals

Neveen Shlayan\textsuperscript{1}, Abdullah Kurkcu\textsuperscript{2}, and Kaan Ozbay\textsuperscript{3}

November 1, 2016

\textsuperscript{1}Assistant Professor, Department of Electrical Engineering, The Cooper Union for the Advancement of Science and Art (e-mail: nshlayan@cooper.edu).

\textsuperscript{2}Ph.D. Student, Civil and Urban Engineering and Center for Urban Science and Progress (CUSP), Tandon School of Engineering, New York University (e-mail: ak4728@nyu.edu).

\textsuperscript{3}Professor, Civil and Urban Engineering and Center for Urban Science and Progress (CUSP), Tandon School of Engineering, New York University (e-mail: kaan.ozbay@nyu.edu).
1. Project Overview

2. Technology Description

3. Data Validation and Filtering

4. Pilot Tests

5. Final Remarks
Projects

- **Real-Time Estimation of Transit Origin-Destination Patterns and Delays Using Low Cost Ubiquitous Advanced Technologies**
  - Funded by: University Transportation Research Center (UTRC)
  - Collaboration with: New York University Polytechnic (NYU Poly) & Rutgers

- **Spectral-based Controllability-preserving Pedestrian Evacuation Network Synthesis Using Multilayered Estimation Models in Real-Time**
  - Funded by: University Transportation Research Center (UTRC)
  - Collaboration with: Center for Urban Science & Progress (CUSP) at NYU Poly
What? Why?

**What are we looking for?**
- Better understanding of pedestrian traffic and extraction of key parameters
  - Wait times
  - Origin-destination (OD) flows

**Why does it matter?**
- Transportation demand forecasting
- Operational evaluation and enhancement of public transportation
- Pedestrian evacuation
- Community wellness
- Safety
Existing Practices

- **Annual travel surveys** are generally used to obtain pedestrian data which are not sufficient.

- **Video-based** pedestrian detection technologies are common; however, they are costly and complex.
  - Multi-person tracking; not scalable [Fan et. al]

- **Multi-sensor systems**, generally not scalable.
  - Infrared cameras, laser scanners, and motion sensors; limited to individual pedestrians [Fardi et. al]
  - LIDAR and a single camera pedestrian detection; can be improved when combined with WiFi and BT [Premebida et. al]

- **Bluetooth detectors**
  - Bluetooth sensors to measure the time for passengers to clear the security screening checkpoint [Bullock et. al]
Objective & Challenges

**Utilizing Bluetooth and WiFi technology to:**
- Establish a general framework through data-driven pedestrian modeling within transit stations
- Obtain time-dependent origin-destination (OD) demands
- Estimate station wait-times of transit bus and subway users
- Estimate average hourly, daily, weekly volume, and delays

**Advantages**
- Cost-effective
- Portable
- Easy to deploy
- Scalable
- Rich data and adequate sample size

**Challenges**
- Lack of sensor infrastructure
- Unrestricted movement of pedestrians
- Filtering, sensor placement, and sensor features algorithms are location and pedestrian nature dependent and not so easy to develop a one fits all solution
Overall System Architecture

Raw data includes: Timestamp, MAC address, BT/WiFi, Received Signal Strength Indication (RSSI)
Hardware

- Android tablet manufactured by ASUS (Nexus 7) with the following specifications:
  - 1.2 GHz CPU, 1 GB memory, 16 GB storage
  - Battery life: 9.5 hours and additional 6-7 hours with external batteries
Software - Traffic Tracker Application

- **Scan**: Start a new scan. The user has to name the new scan such as “Floor 2”. It is possible to get location updates providing GPS locations of a device when there is an internet connection.

- **Database**: After the scan is stopped, the app automatically creates a final table under the “Database” section. It shows the total number of records, scan name, duration, and occurrences of the same devices. These tables are saved in a relational database and can be imported to a text file.

- **Files**: This function allows users to view imported text files.
**Software - Website Application**

- **Anonymization**: The MAC address is double encrypted by deleting part of the MAC address and then encrypting the remaining part.

- **Data Access**: Cloud-based server that connects to all active devices and ensures data transfer between the device and the server.

- **Remote Self-Diagnostics**: Current reporting status of each device, power levels of battery powered devices, and possible data errors.
Double detection: Understand how to deal with a device that is detected by 2 or more sensors at the same time or within a very small amount of time that is less than the estimated transition time.

Radial detection: Determine whether we are able to distinguish between the devices detected outside or inside the building and those that are detected at other floors.

Understand how RSSI values are related to various distance/speed/devices.
Controlled Experiments

Controlled Experiments:
1. Path Tracking
2. Counting

Start at a distance slightly greater than the detection range
The paths were identified with a good accuracy.

The range of the RSSI the sensors picked up are somewhat consistent for all three devices. RSSI between -70 and -80 gives the best results. It is about 10-20 feet from the sensor.

The same MAC address is detected at multiple sensors at once.
The presence of devices that are not necessarily associated with pedestrians within the detection perimeter was an issue.

The detection range cannot be directed and is radial in nature.

There are some errors associated with the sensors accurately reporting the location of the devices.
Addressing Some Issues

- **Location specific filter** is developed to eliminate data entries that may have originated as a result of the aforementioned scenarios.

- **Sensor placement strategy** is needed. One that takes into account the shortest distance between sensors (particularly when placed on different floors), the detection range, and system observability.

- **RSSI levels** in combination with before and after data can be used to filter out and correct false alarms.

- **Sensor fusion** mechanisms, i.e. infrared, may be utilized for data validation.
The initial objectives were to test the devices and to examine the data keeping in mind the estimation of OD flows and wait times for public transit. Multiple pilot tests were conducted in NYC; however, we will discuss only two studies:

- The Atlantic Avenue Subway Station Test, and
- The Port Authority Transit facility in Manhattan
Atlantic Avenue Subway Station

Three different tests were conducted at this location encompassing various periods.

**Table:** Period of data collection by device.

<table>
<thead>
<tr>
<th>Tablet 1</th>
<th>Tablet 2</th>
<th>Tablet 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/30 - 8/3</td>
<td>7/30 - 7/30</td>
<td>7/30 - 8/3</td>
</tr>
<tr>
<td>10AM - 9AM</td>
<td>10AM - 4PM</td>
<td>10AM - 10AM</td>
</tr>
<tr>
<td>-</td>
<td>8/13 - 8/14</td>
<td>8/13 - 8/13</td>
</tr>
<tr>
<td>11AM - 11AM</td>
<td>11AM - 6PM</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>8/14 - 8/18</td>
<td>8/14 - 8/16</td>
</tr>
<tr>
<td>11AM - 6PM</td>
<td>11AM - 1AM</td>
<td></td>
</tr>
</tbody>
</table>
### Table: Total Number of Detected Devices and Average Waiting Times.

<table>
<thead>
<tr>
<th># of recorded devices</th>
<th>Average Wait (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tablet 1</td>
</tr>
<tr>
<td>Tablet 1</td>
<td>2762</td>
</tr>
<tr>
<td>Tablet 2</td>
<td>1221</td>
</tr>
<tr>
<td>Tablet 3</td>
<td>5133</td>
</tr>
</tbody>
</table>

![Graph showing % Movement over Time]
Key Observations

1. **Sample size:** A relatively large number of BT enabled devices were detected by all three tablets.

2. **Basic Statistics:** Average waiting times at platform 3 (Tablet 3) increase from 2.7 minutes to 3.02 minutes after August 14th. A slight increase in waiting times at platform 2 (Tablet 2) after August 14th is also observed.

3. **Irregular behavior:** We saw an increase in movement percentages at platform 3 (Tablet 3) after August 14th.
Port Authority Transit Facility

The sensors are placed in such a way that focuses on the movements from two entrances, E1 and E2, to four different gates, G1, G2, G3, and G4, and attempts to find correlations or clear patterns. It also investigates the potential data discrepancies and sensor malfunctions.
### Results

**Table:** Counts summary of E1 and E2 entrances.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1-WiFi</td>
<td>34</td>
<td>1803</td>
<td>444.1</td>
</tr>
<tr>
<td>E1-BT</td>
<td>1</td>
<td>77</td>
<td>18.5</td>
</tr>
<tr>
<td>E2-WiFi</td>
<td>48</td>
<td>5212</td>
<td>1935</td>
</tr>
<tr>
<td>E2-BT</td>
<td>1</td>
<td>107</td>
<td>22.35</td>
</tr>
</tbody>
</table>

**Figure:** Number of detections for a week at the E1 entrance.
Calculating accurate waiting times requires extensive filtering since it is possible to have some recurring MAC addresses.

Mapping the regular users of the bus line might help reduce some errors.

Relaxing the signal strength filter, which detects only the individuals who are really close to the sensor at the moment, might help to accurately find the initial arrival time to the gate.
Case-specific models have to distinguish between system variations. We are currently developing stochastic Markovian-based models that will enable accurate depiction of a transit user process.

We are able to extract key parameters from the data such as the flow and transition matrices to feed the Markovian process.

This allows us to obtain indicative properties and quantities of the system such as convergence, time to absorption, absorption probabilities, and state density distributions.
Preliminary Results

Transition matrix, $T^n$ absorbed as $n \to \infty$

![Transition Probability Evolution](image)

Density Distribution at Various Locations

![Density Distribution](image)
Conclusions and Future Work

- **Rich data:** The preliminary examination of the BT and WiFi data shows a great promise in its usefulness for evaluation and enhancement of public transportation.

- **Variable nature:** We have established, through examining the data collected from our pilot tests, that the nature of pedestrian movement varies with respect to the public transportation system at hand (i.e. a bus terminal versus a subway station).

- **Modeling:** We are currently looking at stochastic models that are specific to pedestrians in order to be able to estimate and predict OD flows and wait times based on demand.

- **Filtering:** Location and system specific development and design of filtering techniques, sensor placement algorithms, and sensor features have to be developed for more reliable modeling.