COST-EFFECTIVE SENSOR NETWORK TOPOLOGY FOR UBIQUITOUS BLUETOOTH READER DEPLOYMENT IN URBAN NETWORKS

Isaac Kumar Isukapati
isaack@cs.cmu.edu

Gregory J Barlow
gjb@cmu.edu

Stephen F Smith
sfs@cs.cmu.edu

Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213 USA

Submission Date: August 1, 2014
Word count: 4,333 + (7 x 250) = 6,083
ABSTRACT
Travel time, an important performance measure for transportation systems, has traditionally been studied indirectly, but new technologies have made it possible to observe travel times directly. An increasingly popular method for travel time estimation is the use of a network of Bluetooth MAC address readers, where sampled addresses can be matched and travel times estimated. While most studies have addressed deployment for single roads, usually freeways or large arterials, the ubiquitous deployment of Bluetooth readers in a dense urban network raises new questions about network topology. In this paper, we explore the pros and cons associated with node (intersection) versus monument (mid-block) deployment of readers for dense, urban networks. A low-cost Bluetooth reader design developed for these experiments is described, and we present findings from four different deployments using these readers. We conclude that locating MAC readers at nodes is easier and more cost effective than deploying at monuments (where previous studies have recommended that readers should be located), without loss of data quality. The results in this paper show that MAC readers located at nodes are able to capture turning movements effectively, and in many cases outperform readers located at monuments.
1. INTRODUCTION

Travel time information is an important indicator of transportation system performance \((2, 5, 6)\). Traditionally, travel times have been studied “indirectly” from field-based observations of counts and occupancy transformed into spot speeds or observed in snippets via license plate surveys and other techniques. The advent of AVI (Automatic Vehicle Identification) and AVL (Automatic Vehicle Location) technologies has made it possible to observe travel times directly. Automated Vehicle Identification (AVI) data collection sources, which include Bluetooth readers, detect a passing vehicle at one sensor, then re-identify the vehicle at a second sensor, allowing the vehicle’s trip time between two points to be directly computed.

Wasson et al. \((13)\) were first to suggest using Media Access Control (MAC) ID matches to estimate travel times. Since then, the area of focus of most studies that have used Bluetooth MAC reader technology to estimate travel times, has been freeways or large arterials. Such settings require only a sparse deployment of readers on an ad hoc basis to effectively capture the movement of most vehicles. For example, Quayle et al., \((9)\) used Bluetooth data as a surrogate method to floating car studies. Stevanovic et al. \((12)\) focused on testing the reliability of MAC readers to measure travel times. Other similar research efforts \((11)\) have endeavored to extract/predict travel time information from Bluetooth data. In an ideal case, one should collect travel time data continually on every link in the network to improve the accuracy of travel time prediction. This capability becomes increasingly important in the case of urban grid networks, our focus in this paper, where vehicle paths through the network are much more diverse. However, according to a study conducted by Cambridge systematics \((4)\), the cost of installing a single commercial Bluetooth reader is about $10,000. In this sense ubiquitous deployment of MAC readers is not a viable option.

The question is: what should the network topology be from the standpoint of reader deployment if the goal is to ubiquitously deploy these readers permanently in dense urban arterial networks. In this regard, Asudegi and Haghani \((1)\) provide guidance on optimal number and location of Bluetooth readers. SHRP2 L-02 suggests placing sensors at monuments (mid-blocks on links) as opposed to nodes (at intersections) \((7)\). The question about node versus monument is an important one for the following reasons: 1) the total number of readers deployed in the network is a function of where they are deployed (nodes require fewer readers than monuments); 2) MAC readers located at nodes can easily be powered through traffic controller cabinets, whereas in situations where power is not available at a monument, large capacity batteries and possibly solar panels are necessary to power readers for long-term deployments; 3) readers located at nodes may be connected via existing networks used for traffic signal control, whereas the readers located at monuments would likely require expensive dedicated communications, such as cellular modems for each reader. If MAC readers are sufficiently close together, wireless mesh networking could reduce the number of cellular modems required, though the reliability of this approach is questionable. Networking MAC readers not only allows easy data retrieval, it also provides an effective mechanism to synchronize the clocks of all readers. In cases where neither local networking nor remote communication is feasible, clocks could be kept in sync via GPS. For these reasons deploying readers at nodes would likely incur lower deployment costs than deploying them at monuments.
In this paper, we explore the pros and cons associated with node versus monument deployment of readers. We present findings from pilot studies conducted at four locations in Pittsburgh, Pennsylvania. In contrast to the prior recommendations of SHRP2 L-02 (7), our results suggest that placement of readers at intersections may be more cost-effective in the case of urban networks, with little impact on data quality.

2. HARDWARE DESIGN GOALS

Since the goal of this paper is to answer the questions associated with ubiquitous deployment of MAC readers in dense urban networks, custom low-cost Bluetooth MAC readers were designed and built for this research. These MAC readers are ideal for flexible experimental data collection and serve as an early prototype for more permanently installed readers that may be designed in the future. Though not identical to commercial MAC readers, there are enough similarities that this research should be applicable to practical applications using commercial readers.

The basic parts of a Bluetooth MAC reader are a Bluetooth radio with antenna, a computer to capture and store (or retransmit) MAC addresses, a power supply, and a housing for these components. The specific implementation can vary widely, from permanently installed readers in existing cabinets that send data directly to servers for processing to small, portable readers deployed for short periods where data is processed once readers are pulled from the field.

Several considerations drove the design of the MAC readers used to collect the data for this paper. First, readers should be composed of commodity parts that could be easily acquired in small volumes at a low enough cost that high-density deployment would be feasible in the future. The readers should be small, portable, and easy to mount in a variety of locations without relying on existing cabinets for installation. Readers should have at least 24 hours of battery life, with the potential for permanently installed versions in the future.

3. LOW COST MAC READER CONSTRUCTION

The prototype reader used in this work is based around the Raspberry Pi Model B single-board computer. The widespread availability, low cost, ease of use, and computational power of the Raspberry Pi make it an ideal experimental platform. Low cost USB Class 1 micro Bluetooth radios with integrated antenna were used; a wide variety of suitable radios are currently on the market. A Secure Digital (SD) card is required to boot the operating system and to store data, 8 GB cards were used for these experiments. A 12 volt, 8 ampere-hour sealed lead-acid battery provides power, though a higher capacity battery would allow longer deployments. A universal battery elimination circuit (UBEC) switch-mode DC regulator converts the 12 volts from the battery to the 5 volts required by the computer. A weatherproof outdoor enclosure holds the entire MAC reader assembly. The case used for this prototype was 9.45 x 6.30 x 3.60 inches (240 x 160 x 91.44 mm) and designed to IP65 of IEC 529 and NEMA 1, 2, 4, 4x, 12 and 13 specifications. A larger battery would require a larger case. The total cost of all components for each of these MAC readers is less than $120. A picture of one of the MAC readers is shown in Figure 1.

The prototype MAC readers run Pidora, a Fedora-based Linux distribution designed for the Raspberry Pi. The Bluez Linux Bluetooth protocol stack implements the Bluetooth wireless
standards specification. PyBluez provides wrappers around system Bluetooth resources for use in Python. The scanning software for these experiments, written in Python, is based on PyBluez example code for inquiry with received signal strength indication (RSSI).

Figure 1: A picture of MAC reader built for this research study

These low cost MAC readers are effective for collecting short-term data, but extensions and improvements would be necessary for longer-term deployments. An improved reader design would need to address three main issues: how to power readers, how to retrieve data, and how to keep the clocks on readers synchronized. Ideally, readers could always be permanently installed in locations where readers could be powered by the electrical grid and connected to a network for data retrieval and time synchronization, but this isn’t always possible. For MAC readers located at monuments, where it might not be possible to connect to either mains power or an existing network, larger batteries could be deployed, and communications could take place using a cellular modem or a wireless mesh network if readers are sufficiently close together. Locating readers at the nodes, in or near a traffic controller cabinet, would likely be the simplest solution in most cases, given the easy access to power and possible network access. As has been done in many studies (9, 12), a reader could be located within the cabinet with an external antenna mounted on top. This might limit the effectiveness of the reader, since cabinets are not always sufficiently high above the road surface, longer antenna cables reduce the range of the reader, and external antennas may be prone to vandalism. It may be more effective to install a self-
Isukapati, Barlow, Smith

contained reader near the cabinet with a power and networking tether. Depending on the climate and how the MAC readers will be used, the Raspberry Pi could be replaced with a similar single-board computer with a wider operating temperature range. Though this would increase the price of each MAC reader, the total cost could still be kept under $250, well below the price of a commercial reader.

4. FIELD DATA COLLECTION

Experiments were conducted and data was collected at four different locations in the east end of Pittsburgh. Below is a brief description of each site and its intended role in our study.

1) Baum - Centre loop: The purpose of this placement is to check the percentage of times that known MAC addresses are captured by fixed readers; Figure 2(a) shows the location of reader deployment at this site.

2) Centre and Millvale: The purpose of this placement is to see similarities/differences between node and monument data collected at a single intersection; data was collected at this location from July 7-8, 2014 for a period of 24 hours. Figure 2(b) shows where five MAC readers were stationed. Readers 3 & 4 are deployed at the node; reader 5 on Millvale and readers 1 & 2 on Centre Avenue.

3) Centre Avenue: The purpose of this placement is to extend the findings from the previous single intersection deployment to a small network; again data was collected for a period of 24 hours spanning July 16-17, 2014. Figure 2(c) shows the deployment of six readers. Here, readers 2, 5 & 6 are stationed at mid-block monuments whereas readers 1, 3 & 4 are stationed at nodes.

4) Baum Boulevard: The objective of this placement is to address node versus monument deployment for a small but busy arterial network; data was collected for a period of 24 hours spanning July 21-22, 2014. Figure 2(d) shows the locations of 11 readers; readers 1, 5, 6, 7 & 9 are deployed at monuments whereas readers 2, 3, 4, 8 & 11 are deployed at the nodes.
Figure 2(a): Schematic of reader deployment on Baum – Centre loop

Figure 2(b): Schematic of reader deployment on Centre and Millvale
4.1. Mounting the readers and data retrieval

Since the experiments described in this paper were performed at a variety of locations and for short periods of time (24 hours), all mounting was temporary. Readers were mounted approximately eight feet above the road surface (as recommended in (3)). To mount a MAC reader, two 48-inch zip ties were secured approximately 10 inches apart on an existing pole near the side of the road (typically a streetlight or traffic signal pole). Four 8-inch zip ties were
attached to the 48-inch zip ties and flanges at the four corners of the MAC reader enclosure. The battery was always mounted at the top of the enclosure.

Readers are deployed in the field using a common procedure. Since the Raspberry Pi lacks a real-time clock, readers are connected to the network at boot time, and then the clock is set using the network time protocol (NTP). The USB Bluetooth device is turned on, and then continually scans for discoverable devices, capturing the MAC address and RSSI for each device along with the time of detection. All entries are recorded for processing. After initialization, MAC readers are disconnected from the network, enclosures secured, and installed in the field, typically for 24 hours. Readers are then brought back to the lab to retrieve data.

**4.2. Note on discoverable Bluetooth devices**

Bluetooth MAC readers constantly scan for discoverable Bluetooth devices. Even though a wide variety of Bluetooth devices exist and would provide usable travel time information, not all devices will be detected by these MAC readers, as devices must be in discoverable mode to respond to the types of scans MAC readers perform. Many late model cars are now equipped with Bluetooth, and likely comprise the largest set of Bluetooth devices usable for travel time estimation. While most mobile phones are equipped with Bluetooth, most of those phones are no longer discoverable by default, even when a phone is actively using Bluetooth. Many Bluetooth accessories that might be located in vehicles, such as hands-free headsets, are discoverable. The discoverability of Bluetooth devices varies by manufacturer, but enough of these devices are detectable to make MAC readers effective for estimating travel times.

**4.3. Ascribing passage times**

The first analysis task is to assign a passage time to each unique vehicle trip; passage time indicates when a given vehicle passed a specific reader. Here each raw observation includes: the MAC address, signal strength, and UNIX time stamp. Figure 3 describes pseudo code for ascribing passage times. The detection with the highest signal strength is used for ascribing passage times as in (10). The methodology described in (8) was used to filter outliers in the dataset, and we assume that if two observations of a given MAC ID are separated in time by more than 2 minutes that they represent two distinct trips.
Input:

\[ D = \begin{bmatrix}
I_1 &=& [o_{i1}, o_{i2}, \ldots, o_{in}]
I_2 &=& [o_{i21}, o_{i22}, \ldots, o_{i2n}]
\vdots
I_m &=& [o_{im1}, o_{im2}, \ldots, o_{imn}]
\end{bmatrix} \] (set of observations associated with ‘m’ unique MACIDs)

Where \( o_y = (t_y, S_y) \)

Output:

\[ T = \begin{bmatrix}
I_1 &=& [t_{i1}, t_{i2}, \ldots, t_{in}]
I_2 &=& [t_{i21}, t_{i22}, \ldots, t_{i2n}]
\vdots
I_m &=& [t_{im1}, t_{im2}, \ldots, t_{imn}]
\end{bmatrix} \] (time stamps associated with each unique trip)

Procedure for clustering and ascribing passage times

For each \( I_i \) in \( D \)

\[ O = [o_{i1}, o_{i2}, \ldots, o_{in}] \]

Sort observations in ‘O’ in chronological order (oldest to newest)

Create a new set of passage times associated with \( I_i \)

Start record = \( o_{i1} \)

For \( o_j \) in \( O \):

If \( \Delta t_{(i,j)} > 2 \) minutes:

In partition set from start record up to record ‘j’,

a) find the record with highest signal strength (if ties, choose the earliest)

b) treat the time stamp associated with that record as the passage time for the trip

c) add passage time to the set of passage times associated with \( I_i \)

Update start record = \( o_{j+1} \)

Add all passage times associated with \( I_j \) to \( T \)

Return \( T \)

---

**Figure 3: Procedure for ascribing passage times**

---

### 5. EVALUATION OF RESULTS

As indicated earlier, our ultimate interest is in understanding the performance differences of deploying MAC readers at nodes as opposed to monuments. Is it the case that readers positioned at nodes are able to capture most if not all MAC addresses associated with various turning movements? Or, as previous analysis has suggested, are monument placements necessary to provide adequate coverage? The analysis presented in this section is targeted toward answering these questions.
5.1. Known MAC address capture

Though MAC readers may detect many unique Bluetooth devices, it is important to determine, given a known Bluetooth device, the likelihood of a reader detecting that device. If the likelihood of detecting any given device is low, then the usefulness of Bluetooth travel time measurements is limited, even with high Bluetooth penetration. Many different factors affect this likelihood, including vehicle speed, distance between the reader and the device, occlusion by other vehicles, location and class of the Bluetooth device within a vehicle, and the construction of the MAC reader itself. To ensure that the experimental readers developed for this study are sufficiently likely to detect a given device, we inserted Bluetooth devices with known MAC addresses into the test vehicles that were used to carry out the first experiment.

More specifically, seven readers were deployed in the Baum-Centre loop network shown in Figure 2(a) and three Bluetooth devices with known MAC addresses were driven clockwise through the network in two vehicles. MAC readers were located both at nodes (readers 1, 4, 8, 10) and monuments (readers 2, 7, 9). The devices located in vehicles (readers 3, 5, 11) were identically constructed MAC readers where each Class 1 Bluetooth device was placed in discoverable mode.

Several readers were placed in less than ideal positions, some intentionally, some unintentionally. Reader 4 was placed directly above the cabinet at Centre and Morewood, which does not have line of sight to the westbound queue on Centre. The test vehicles were often occluded from reader 1 during the right hand turn at Centre and, since there are effectively two lanes at Centre and Millvale. A boom truck occluded reader 10 for the entire experiment at Baum and Millvale (the reader had been deployed the afternoon before the experiment). The percentage of trips where each fixed reader detected each known Bluetooth device is shown in Table 1. Additionally, an Android smartphone was put into discoverable mode for the first circuit of the loop, and was detected by every fixed reader. Based on these results, detection rates were high enough (even for the occluded readers) to suggest that the MAC readers used in this study are sufficiently likely to detect vehicles carrying discoverable Bluetooth devices. For the remaining experiments, results were obtained using actual traffic flows.

<table>
<thead>
<tr>
<th>Known_ID</th>
<th>Location</th>
<th>% of times known_Id was seen</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Passenger Seat car -1</td>
<td>90 100 80 100 100 100 80</td>
</tr>
<tr>
<td>11</td>
<td>dash board car -1</td>
<td>80 90 90 90 100 90 60</td>
</tr>
<tr>
<td>5</td>
<td>Passenger Seat car -2</td>
<td>44 100 78 89 89 78 44</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics of known Id captures by various readers

5.2. Scenarios requiring two readers at intersection

To answer the question of whether a single MAC reader at a node is sufficient, a second experiment was performed where data was collected around a single intersection (Centre and Millvale), shown in Figure 2(b). Readers 1, 2 & 5 were deployed at monuments while readers 3 & 4 were deployed at the node, with reader 3 directly above the traffic controller cabinet very near the road, and reader 4 set back further, diagonally across the intersection. Table 2 presents summary statistics for Origin – Destination (O-D) pairs (1, 2), (1, 5) and (2, 5). There are five
columns in this table: the first two columns represent the O-D pair (both directions); the third and fourth columns give the percentage of O-D trips observed by readers 3 and 4 respectively; and the values presented in column 5 represent the number of unique MAC addresses captured by both origin and destination. The table suggests the readers at the node are capturing a very high percentage of the O-D trips with one exception (reader 4 captured only 83% of the (1, 2) trips). The possible explanation for this could be the combination of location of reader 4 and the occlusion due to left turns from Centre onto Millvale. More generally, we conjecture that very wide intersections may require two readers at the node to capture all turning movements, but that a single reader will typically be sufficient. It is worth exploring this issue further in future research.

<table>
<thead>
<tr>
<th>O - D</th>
<th>% observed at the node</th>
<th>No of obs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>94</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>97</td>
</tr>
</tbody>
</table>

**Table 2: Summary statistics for location # 1**

### 5.3. Comparisons of readings at Nodes versus Monument

Data collected at Location #3 (Centre Avenue, shown in Figure 2(c)) and Location #4 (Baum Boulevard, shown in Figure 2(d)) were used to examine the tradeoff between node and monument placements. Let’s first consider the Baum Boulevard case. Six O-D pairs on Baum Boulevard were selected to highlight differences in node versus monument data: O-D pair (6, 10) represents through movements starting or ending west of Morewood on Baum (at reader 6) and starting or ending east of Liberty on Baum (at reader 10). Though readers 6 & 10 are monuments, both can be considered functionally equivalent to nodes in the sense the vehicles travel at lower speeds due to queuing at nearby intersections. The placement of these readers was intentional to highlight differences in reader placement. Every (6, 10) O-D trip passes through 3 nodes (readers 11, 3, & 2) and 2 monuments (readers 1 & 5). Similarly, a second O-D pair (6, 9) is a right turn movement from Baum on Cypress or a left from Cypress to Baum. Every vehicle that makes either (6, 9) trip passes through two nodes (readers 11 & 3) and one monument (reader 1). O-D pair (6, 8) is again a right turn movement from Baum onto Liberty or a left from Liberty to Baum. These trips pass through three nodes (readers 11, 3 & 2) and two monuments (readers 1 & 5). O-D pairs (6, 7) represent either a left turn movement from Baum onto Liberty or a right from Liberty to Baum. These trips go through the same nodes and monuments as (6, 8). Similarly, every vehicle that makes (9, 10) or (9, 7) trips passes through two nodes (reader 3 & 2) and one monument (reader 5).

Table 3 presents summary statistics for the six O-D pairs described above. This table has eight columns: the first two columns present information on the choice of the O-D pair; the values presented in column 8 represent the number of unique MAC addresses captured by both origin and destination. Values in columns 3 – 7 represent the percentage of O-D trips captured by various readers at nodes/monuments en-route. Please notice the percentage of trips captured at monuments (readers 1 & 5) are always lower than those captured at nodes (readers 2, 3 & 11).
Table 3: Summary statistics for location # 4

Similar analysis was conducted on data collected at location #3; Table 4 summarizes the results. Here, O-D pair (5, 6) includes trips between Millvale and Neville. Similarly, O-D pair (2, 6) includes left turn movements from Centre onto Neville and right turn movements from Neville onto Centre. O-D pair (2, 3) represents vehicle through movements on Centre Devonshire and Melwood. Each of these O-D trips passes through two nodes (readers 1 & 4).

Table 4: Summary statistics for location # 2

6. CONCLUSIONS AND RECOMMENDATIONS

This paper answers network topology questions concerning the permanent, ubiquitous deployment of MAC readers for traffic performance analysis in urban road networks. In this context, where vehicle traffic flow patterns are typically more diverse than previously considered expressway and arterial network configurations, we explore the pros and cons associated with node versus monument deployment of readers. We present findings from pilot studies conducted on four locations in Pittsburgh, Pennsylvania. Since the goal of this paper is to answer the questions associated with ubiquitous deployment of MAC readers in dense urban networks, custom low-cost Bluetooth MAC readers were designed and built for this research. To ensure that the experimental readers developed for this study are sufficiently likely to detect a given device, three Bluetooth devices with known MAC addresses were first driven repeatedly in two vehicles through a network of seven fixed readers. Based on the results, detection rates were high enough (even for some occluded readers) to suggest that the MAC readers used in this study are sufficiently likely to detect vehicles carrying discoverable Bluetooth devices. Results of the subsequent pilot studies indicate that the percentages of trips captured at monuments are always lower than those captured at nodes. This is encouraging for two reasons: 1) the placement of readers strictly at nodes reduces the total number of readers required to fully instrument the network; 2) placement at nodes is more cost-effective while producing a better overall detection capability. We also conjecture that very wide intersections may require two readers at the node to effectively capture all turning movements, but that a single reader will typically be sufficient.
REFERENCES


