Application Guidance for Bluetooth Travel Time Data Collection

by

Corresponding author:
Edward J. Smaglik
Northern Arizona University
Department of Civil Engineering, Construction Management, and Environmental Engineering
P.O. Box 15600
Flagstaff, AZ 86011
edward.smaglik@nau.edu

Craig A. Roberts
Northern Arizona University
Department of Civil Engineering, Construction Management, and Environmental Engineering
P.O. Box 15600
Flagstaff, AZ 86011
craig.roberts@nau.edu

November 14, 2013

TRB Paper #14-2727

Word Count: 5741 words + (7 x 250 words/Figure-Table) = 7491
Abstract

In our performance measure based society, travel time information is critical for the efficient operation of a transportation system, providing baseline operational information as well as feedback after infrastructure and/or ITS changes are implemented. In the last six years, generation of travel time through the use of Bluetooth MAC address tracking has become a relatively low cost acceptable surrogate for more expensive traditional methods of travel time generation, such as floating car studies or license plate matching. While studies over multiple modes have validated this technology, MAC address tracking is not a panacea for all situations and the limitations of the technology must be considered prior to embarking on a deployment. The objective of this paper is to provide practical guidance to practitioners considering such a deployment, and will do so in three steps. First, a brief literature review is conducted with a focus on items that would be of interest to a practitioner investigating a possible Bluetooth deployment. Second, a case study is presented of a vehicular Bluetooth travel time analysis deployment which was hampered by data insufficiency and software usability; specific focus is given to the data analysis conducted with commercially provided hardware and software. Finally, the paper concludes with a section describing the lessons learned from the case study, with the goal of providing guidance to those considering their own Bluetooth applications.

Introduction, Motivation, and Objective

In the arena of travel time measurement, the method of tracking Bluetooth media access control (MAC) addresses across a transportation facility to infer travel time is relatively young, with the first real work in the area beginning about six years ago (1). However, in that relatively short time period, the technology has been widely deployed for multiple purposes, primarily the generation of travel times for various modes, though Origin-Destination and pedestrian studies have been conducted as well. Much of this work, from the construction and deployment of the units to the development of algorithms to parse collected data has been documented in various industry journals. This, combined with the relative ease of and low cost of installation, has turned the technology into somewhat of a panacea: “Let’s just throw some Bluetooth sensors out there and see what we get.”

In its infancy, the analysis of collected Bluetooth data required building the data collection devices, as well the development of algorithms to parse the data in an attempt to infer useful information from the collected data. The latter part of this task is the more challenging part, as the parts to build a Bluetooth data collection unit are around $500, and the construction fairly simple (transceiver, antenna, power source, data storage, and optional communications). As the method became more mature, multiple commercial plug and play solutions populated the marketplace, some directly derived from initial research work. With the cost of these units well into the thousands of dollars, vendors provide software solutions for data management in exchange for the additional capital required beyond the ‘Do-It-Yourself’ method. Initial researchers in the area had no choice but to develop their own software solutions for data mining and management, and as such ended up with solutions tailored to their unique situation. However, off-the-shelf software provided by a vendor may or may not have the functionality required to properly address the problem at hand.
It is expected that the vast majority of practitioners interested in deploying Bluetooth for travel
time data collection will go the commercial route, as the specialty knowledge required for
developing a homegrown data management solution as well as the time required to complete
such an endeavor are not likely to be readily available at a typical transportation management
agency. This paper is not written to push the envelope of Bluetooth travel time research; its
objective is to provide practical guidance to transportation practitioners considering such a
deployment, and will do so in three steps. First, a brief literature review will be conducted with a
focus on items that would be of interest to a practitioner investigating a possible Bluetooth
deployment. Second, a case study of a vehicular Bluetooth travel time analysis deployment
using commercially purchased equipment will be presented, with specific focus on the data
analysis conducted with the provided software. Finally, the paper will conclude with a section
describing the lessons learned from the case study, with the goal of providing this guidance.

Literature Review

While the generation of travel time through Bluetooth MAC address matching is less than a
decade old, there is much in the literature regarding many aspects of the process, from building
a unit to processing the data. The review which follows will provide a high level view of several
different topics within the literature, focusing on topics that are of import to practitioners
interested in deploying this technology.

General concept and sensor construction

The general concept of this type of travel time data collection is to track specific MAC addresses
at different points within a network, and then infer travel time from the time stamped data. Every
Bluetooth radio, whether it resides within a smartphone, hands-free headset, automobile, or
computer, leaves the manufacturer with a unique MAC, non-traceable MAC address. Using a
second Bluetooth radio mounted along the side of a roadway, these addresses can be
observed, time stamped, and logged. Then, the logs from different data sites can be compared,
and a rough travel time can be inferred (2). The equipment required to build a Bluetooth data
collection device can be obtained for around $500, and includes the following:

• Bluetooth radio, class dependent upon need
  o Class 1, 100 meter range
  o Class 2, 10 meter range
  o Class 3, 1 meter range
• Antenna
• Power source, battery or solar
• Data logging device (field pc, or other computing widget)
• Data storage location
• Communication device (such as a cell modem), if desired
• External case, if necessary

This computing hardware can be mounted within a powered existing enclosure, with an external
antenna. Alternatively, all of the above hardware, including the antenna, can be mounted within
a weather resistant case. Researchers with Purdue University recommend that the antenna be mounted at least eight feet above the finished roadway grade to maximize the number of MAC addresses detected (3) for vehicular traffic. Placement, transceiver strength, and antenna type vary based upon the desired mode to be sampled.

Data analysis: Outlier Determination
While the construction of the data collection device is relatively straightforward, aggregating useful travel time data from the MAC address logs requires some mathematical work and acceptance of inherent errors. Generally, one is looking for the travel time between two points, or an average speed on a specific segment, as a measure to report. For an arbitrary travel segment on an arterial (Point X is the beginning of the segment; Point Y is the end), there are a number of ways that inaccurate data points can find their way into the data, such as

- A vehicle deviating from the expected route within the segment (1)
- MAC address matches from users not of interest, such as vehicles when studying pedestrians, or vice versa
- A MAC address match from a vehicle making multiple trips along a corridor, where the address is detected at Point X on the first trip, and Point Y on a successive trip (4)

In order to generate an accurate travel time value, these outliers must be filtered from the data or their influence mitigated. Different authors have approached this problem with different methods, some more complex than others. One approach is to determine the central tendency of the data, and then remove data points that fall an arbitrary percentile above or below that central tendency (1, 4), while another is to apply a moving standard deviation to the data set, removing points falling outside an arbitrary multiple of the standard deviation (4, 5).

Penetration
One of the principle benefits of using probe data from Bluetooth sensors is that the data provides a better picture of the actual travel time because many more probes can be detected than with legacy travel time data collection techniques. More probes equal more data points, which can provide for a more precise statistical analysis of the data. Different authors have reported that the penetration rate for motor vehicles, that is the percentage of vehicles on a given roadway with a detectable Bluetooth device, can vary from between 2.0%-3.4% on a freeway corridor (4) and from 1.5%-4.0% on an urban arterial (5). The literature does not have guidance on bicycle penetration, and though work has been limited in the pedestrian area, one source showed that 7% of pedestrians in their study carried such devices. (6, 7)

Other Sources of error
One source of error in the data comes from the characteristic that the Bluetooth MAC address sensor is a zonal sensor, detecting all MAC addresses within its range, with no ability to identify where within that range the address was identified. For example, if a Class 1 Bluetooth transceiver is used for data collection device at Point X and Point Y, a MAC address can be read at any point within the theoretical 100m (328ft) range of each radio. In a worst case scenario, the MAC address is read 100m upstream from the true location of Point X, and 100m downstream from the true location of Point Y, resulting in 200m (656ft) of distance error.
Research personnel from Purdue commented that this error would be negligible if relatively large (2-3 mile) segments were used (1), while personnel at the University of Maryland validated this comment with calculations, showing that the maximum possible error in a calculation of speed, given the transceivers in the example above, is 1.5 mph at 45mph, over a 1.5 mile segment (4).

Research personnel at the University of Washington looked into this problem as well, and determined that a larger detection zone is desirable even though it increases the possible margin of detection, as it provides a larger sample size, which they found would reduce random error. Mounting two MAC address sensors at one location, and then aggregating the data from the two sensors, also was shown to improve accuracy (8). In another paper by the University of Washington, research personnel found that although Bluetooth sensors tended to be biased towards slower moving motor vehicles, because they spend more time within range of the sensor, the travel times generated by the Bluetooth sensors in their study were representative of those provided by license plate recognition (9).

Personnel at the University of Akron conducted a detailed study on the impact that effective sensor range and vehicle speed have on the probability that a unique MAC address is detected by two separate scanning radios. This study found that the probability of detection increases on roadways with a lower functional classification, lower speed limits, and during periods of congestion. This paper also provided information to help fine tune a data collection setup based upon vehicle speed and site characteristics (2).

Privacy

Privacy can be a concern, as generating travel times with Bluetooth technology does involve recording MAC addresses of passing devices. Even though each MAC address is unique, and not linked to any specific user account or central clearinghouse (4), users still may have privacy concerns. These concerns can be mitigated by storing any collected MAC addresses for a limited period of time (1), and/or by filtering the collected MAC addresses, such that it is still possible to match devices at different data collection points, but only a portion of the actual address is kept in a database, making it impossible to recreate the actual MAC address at a later point in time (5).

Effectiveness of Bluetooth data

Despite the challenges and concerns listed in the previous paragraphs of this section, travel times generated with Bluetooth MAC address detection have been shown to be an effective method for travel time generation when compared to alternate methods. In the City of Portland, Oregon, Bluetooth readers were deployed along a signalized corridor to measure the change in motor vehicle travel times effected by a signal timing change. The Bluetooth data correlated with concurrent floating car data, though no statistical comparison was undertaken (5). Personnel at the University of Maryland also compared ground truth Bluetooth travel times with actual floating car data, and did not find the Bluetooth data to be significantly different (4). Similar results have been found when compared with data from license plate recognition software (9). With this type of data reliability, Bluetooth data has been effectively used to address a number of different problems, from the generic travel time and speed data already
discussed, to Origin-Destination studies (10), pedestrian travel time analysis (6), and the
analysis of intersection performance (11).

Case Study: City of Flagstaff Snowplay Congestion

Background
Flagstaff, Arizona, is a city of 65,000 people located 120 miles north of Phoenix, at an elevation
of 7,000 feet. Located in the high desert of Arizona in the shadow of the San Francisco peaks,
Flagstaff, a four-season town, receives on average about 100 inches of snow per year. Given
that Flagstaff is one of the few locations in the state to offer snow recreation, in the form of
skiing and snowboarding at the Arizona Snowbowl, and sledding and tubing at several vendor
operated snowplay areas, a large volume of out of town traffic frequents these locations during
the winter season. All of these areas are located north of the city, as shown on the location map
in Figure 1.

Figure 1: Flagstaff location map
The location of a vendor operated snowplay area is shown as Point A in Figure 1, and the junction of the Arizona Snowbowl access road with the main highway, US 180, is shown as Point B. The only route back to town from these two points is US 180, which is a two lane roadway with no viable alternatives until Point C (Point D is the interchange of I-17 & I-40). Given that all of these areas flush out customers around sunset, heavy congestion occurs during peak snowplay periods, such as academic winter break, Martin Luther King (MLK) Day weekend, and President’s Day weekend. Anecdotal evidence recounted travel times in excess of two hours from the Arizona Snowbowl parking lot to Point C.

After Point C, there are several alternate routes available to drivers, however since the majority of drivers are not local, they are greatly underutilized. The City of Flagstaff (CoF) and the Flagstaff Metropolitan Planning Organization (FMPO) enlisted the help of the local police to direct traffic at Point C towards these alternate routes, as well placed trailblazing signs to highlight one alternative route, with dubious success. These routes are shown in Figure 2, with Point A from Figure 1 as a starting point, and Point D as the terminus. Staff at the FMPO believes that much of the traffic is not attempting to leave town during these times of congestion, but alternate routes were directed to Point D as it is a good reference point for all drivers, including those from out of town. The routes in Figure 2 are not numbered sequentially because they are part of a larger pool of alternative routes developed for this study.
Proposed Remedy

Ultimately, FMPO desired to provide real time travel information for alternate routes to snowplay users, to allow them to make individual route choice decisions. The first step in providing this sort of information to the public was to determine if reliable travel time data could be generated in real time along the various routes. Given the variety of routes under consideration, using Bluetooth technology to develop travel times was thought to be the most cost effective method of developing travel times as a ‘net’ of sensors could be cast across the area. However, there was some concern that traffic volumes might be too low to support this type of data collection on some of the alternate routes. Because of this, the use of cell phone data for travel time
generation, either as a standalone method or to supplement Bluetooth data was discussed, but discarded after it was learned that the cell phone provider used to generate travel time data was one that has very poor reception in the area.

Once the technology was decided upon, staff at the FMPO and research personnel at Northern Arizona University (NAU) held several discussions regarding whether or not to employ a homegrown solution, wherein NAU would build the units and develop the data management software, or to go the commercial, plug and play route. Given the desire to collect data in the current winter season (these discussions were occurring in November and early December 2010), it was decided that a commercial solution would be the preferred route, as it would eliminate any hardware / software development issues that could delay deployment. FMPO staff solicited product information from several vendors, asking questions about hardware, software, and cost. After receiving responses from two vendors, FMPO selected a solution from a vendor where two units would be purchased, and fourteen units would be rented for the duration of the data collection. Each unit would be powered by an onboard 12V deep cycle marine battery, and be able to store data locally (via an SD card) as well as transmit the data via cell modem to the vendor’s central office. The vendor would provide a real-time web interface for viewing the data, as well as a standalone desktop software for post-processing.

**Unit Deployment**

The units were shipped to FMPO early in the week of January 10th, 2011. The vendor’s staff met with CoF, FMPO, and NAU personnel on Wednesday, January 12th, 2011 to discuss the logistics for installing the units, as well as the placement of the sixteen data collection units to generate travel times along nine proposed alternate routes, six additional to those shown in Figure 2. The location of these units across the Flagstaff network is shown in Figure 3. The numbers in the table within Figure 3 describe the specific location of stations, which are also shown on the map.
<table>
<thead>
<tr>
<th>Sensor (Station) Number</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wing Mountain Snowplay Area</td>
</tr>
<tr>
<td>2</td>
<td>US180 S. of Snowbowl Rd</td>
</tr>
<tr>
<td>3</td>
<td>Beale &amp; US180 (N. of Forest)</td>
</tr>
<tr>
<td>4</td>
<td>Sullivan &amp; US180 (S. of Columbus)</td>
</tr>
<tr>
<td>5</td>
<td>US66 W. of Humphreys</td>
</tr>
<tr>
<td>6</td>
<td>Milton S. of Butler</td>
</tr>
<tr>
<td>7</td>
<td>Milton S. of US66</td>
</tr>
<tr>
<td>8</td>
<td>Milton S. of Forest Meadows</td>
</tr>
<tr>
<td>9</td>
<td>I-17 S. of I-40</td>
</tr>
<tr>
<td>10</td>
<td>Beaver S. of Forest</td>
</tr>
<tr>
<td>11</td>
<td>Forest W. of Turquoise</td>
</tr>
<tr>
<td>12</td>
<td>Beaver &amp; Elm (S. of Columbus)</td>
</tr>
<tr>
<td>13</td>
<td>San Francisco S. of Forest</td>
</tr>
<tr>
<td>14</td>
<td>Switzer Canyon N. of US66</td>
</tr>
<tr>
<td>15</td>
<td>Ponderosa Parkway &amp; US66</td>
</tr>
<tr>
<td>16</td>
<td>Butler E. of Ponderosa Parkway</td>
</tr>
</tbody>
</table>

Figure 3: Data collection unit locations

1. The units were deployed on January 14th, 2011, in time to collect data from the upcoming Martin Luther King (MLK) weekend (January 15th – 18th, 2011). The expected life each of battery was
about 14 days, however after about 9 days, some of the units began to lose power. By about 12 days, all of the units were dead. The original plan was to do one battery swap. Because President’s Day weekend (February 18th-20th, 2011) was of import for data collection, staff held off on replacing the batteries in all units until just before that weekend. The units were pulled from service on February 22nd, 2011. The vendor was unsure why the battery life was shortened. FMPO and NAU staff felt it might be from the cell modem, as the service provider used by the vendor, AT&T, does not provide very good coverage in the Flagstaff area, causing the unit to constantly search for service, resulting in reduced battery life.

**Data Analysis: Web Portal**

One of the software products provided by the vendor was a web portal for real time viewing of the data. For the purpose of this work, the following definitions apply:

- **Station**: Individual Bluetooth MAC Address sensor
- **Segment**: A unique directional combination of two stations.
- **Route**: A collection of adjoining segments.

NAU and FMPO staff provided route data for the main route from the snowplay area to the freeway interchange (Point A to Point D in Figure 1), and vendor staff programmed the software to produce the real time travel plot shown in Figure 4.

![Calculated Travel Times](image-url)

**Figure 4**: Real time travel plot
Each color shown in Figure 4 corresponds to the travel time on a segment of the route. This snapshot, take on January 16th, 2011, the Saturday of MLK weekend, shows that there is congestion on the route, with travel time increasing to a peak of about 45 minutes around 6:00 PM, 25 minutes greater than the uncongested travel time of 22 minutes (Figure 4 does not accurately represent the uncongested travel time). Another item to note in Figure 4 is the ‘break lines’ that appear on the plot. An example of this is circled in Figure 4. This occurs when one of the segments does not have sufficient data to generate a travel time estimate; it reverts to a pre-programmed value. The horizontal arrow in Figure 4 highlights the duration of time on this plot where there was enough data to produce a travel time on each of the nine segments of the route: about 4.5 hours. The web portal had the capability to show detailed information for each station and each segment, as well as for routes. In addition, it could have been programmed to show this type of plot for each of the alternate routes, as we, however NAU and FMPO only requested this for one route, as there was no way to compare route travel times with this tool.

**Data Analysis: Desktop Software**

One of the goals of this Bluetooth deployment was to compare travel times concurrently along alternate routes. The vendor provided a post-processing software as a tool to complete this comparison. Data was retrieved daily by the vendor from each station via cell modem, and then forwarded to NAU personnel via an email drop, one email per sensor. Once placed in the proper directory on the computer, the post-processing software could access the data.

**Data Processing**

The first step with this software is to process the data from each station. Once the data from all stations is processed, the user must identify all possible segments to be used in the analysis. For the analysis conducted here, this included each segment in each of the nine proposed alternate routes. With segments identified, the user moves on to processing of the segments. At this point, the software provides several options for the processing.

The first choice controls the period of time in which a MAC address identified at both stations is considered a valid match. For example, if a MAC ID is recorded at Station X at time 12:00 PM, and Station Y at 12:31 PM, it would not be a match if the maximum time period, or window, is set for less than 31 minutes; it would therefore be discarded. The software does provide several default window settings, termed FREEWAY and ARTERIAL. These settings use default maximum windows of 30 and 60 minutes respectively. If FREEWAY or ARTERIAL is chosen, the software compares the default maximum search window with the value of 6 times the actual free flow travel time, and uses the greater value. The software uses free flow travel speeds of 30 mph for arterials and 65 mph for freeways. These values are not user adjustable. For this reason, research staff set the windows manually. This allowed the staff to tailor the window size for each segment, as the length of segments varied from 7 miles to ¼ mile.

The second choice is whether or not to apply a filter, termed ‘IRQ4,’ to remove outlier points from the data. While the details of how this filter is applied is fully described in the software user manual (12), it effectively marks all data points greater than 3 standard deviations from the mean as outliers and rejects them. The closest 30 data points (temporally) to the point in question are used to calculate the standard deviation. This is an important piece of information.
regarding the analysis of routes because the amount of time it will take to generate 30 data points can vary greatly from segment to segment based upon the characteristics of each specific segment, e.g., traffic volume. If traffic volumes vary greatly from segment to segment, the outlier determination time period can differ greatly between segments. For example, Figure 5 shows a 24-hr plot of MAC address matches on two different segments. Each bar in the plot represents the number of matches on the segment during a five minute bin. Figure 5a shows the matches for the Segment 1:2, a two-lane, rural road. Figure 5b shows the matches for Segment 5:6, a 4-lane, urban arterial.

Visual inspection of the plots shows that the Segment 5:6 has many more matches than Segment 1:2. Delving further into the data, the closest 30 data points to a vehicle matched at 5:00 PM on Segment 1:2 span a time period of 2 hours and 35 minutes, from 3:45 PM to 6:20 PM. Conversely, the time period required to span the closest 30 data points on Segment 5:6 is only 30 minutes, from 4:45 PM to 5:15 PM. The vendor believed that the impact of this was minimal, but was not able to provide a statistical analysis to support this conclusion.

**Route Analysis**

With all the segments processed, the calculation of travel times on the various routes was the next step in the analysis. The determination of route travel time required several separate steps. First, the user must calculate summary statistics for each segment to be used in the route analysis, choosing from a bin size of 5, 15, or 60 minutes. For analysis of the data in this project, a 15 minute bin was chosen, given the low volumes in the network. Second, a corridor report for each route was created and exported to Excel. The user must choose one of the
following for the summary statistic in the corridor report (if more than one summary statistic is desired, the report must be run again for each statistic):

- Mean Travel Time
- Median Travel Time
- Number of Samples
- Standard Deviation
- Mean Speed

Table 1 shows the corridor report for Route 1, the main route from Point A to Point D, for both the AM and PM analysis periods used in this work. The steps described above were followed again to generate the corridor report for each additional route. Any comparison between routes needed to be done manually; the software did not support this function.

**Data Sufficiency**

Each data cell in the Table 1 shows the average travel time across a segment during a specific time period. For example, for Segment 2:3 from 9:15 AM – 9:30 AM, the average travel time was 7.66 minutes. This value is the average of all non-outlier data points (MAC address matches) collected during that 15 minute period. A MAC address match is temporally classified by the time logged at the second station of the segment. In this example, that would be the time the address passed station 3. To provide the travel time for an entire route, the software sums the individual mean travel times for each segment for that time interval.

Table 1 highlights several issues with data insufficiency. First, in Table 1a, there are a handful of blank cells, with the majority of them on the lower volume segments, 1:2 and 2:3. These blank cells results in no value for the route travel time. Due to empty data cells, only 3 of the 13 time bins shown in this example provide a travel time for the entire route.

The second issue created by a lack of sufficient data is the inability of the software to determine the central tendency of the desired data stream. Table 1b shows data output for Route 1 from January 15th, 2011, during a PM time period. Looking at the time period of 6:30 PM – 7:45 PM on segment 7:8, a one mile segment with many food and gas stations, the travel time for the segment varies from a low of 2.10 minutes to a high of 52.79 minutes. Given that the travel times for the upstream segment (6:7) are relatively consistent during this time period, it is highly unlikely that the larger values are accurately reflecting the travel time on the segment. It is believed that this variance is due to an insufficient amount of data points around the true central tendency of the data, combined with many users detouring from their trips to stop for food or gas on this corridor. Because of this, larger values are not filtered out by the outlier algorithm, resulting in reported travel times that are most likely much higher than actual.
<table>
<thead>
<tr>
<th>Segment</th>
<th>1:2</th>
<th>2:3</th>
<th>3:4</th>
<th>4:5</th>
<th>5:6</th>
<th>6:7</th>
<th>7:8</th>
<th>8:9</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Begin Time</strong></td>
<td><strong>End Time</strong></td>
<td>0.97</td>
<td>1.46</td>
<td>1.09</td>
<td>0.60</td>
<td>16.51</td>
<td>1.15</td>
<td>16.51</td>
<td>1.15</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>9:15 AM</td>
<td>7.66</td>
<td>0.89</td>
<td>2.66</td>
<td>1.05</td>
<td>0.54</td>
<td>7.61</td>
<td>1.29</td>
<td>7.61</td>
</tr>
<tr>
<td>9:45 AM</td>
<td>10:00 AM</td>
<td>8.13</td>
<td>0.87</td>
<td>1.42</td>
<td>1.03</td>
<td>0.59</td>
<td>13.29</td>
<td>1.22</td>
<td>13.29</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>10:15 AM</td>
<td>0.97</td>
<td>1.25</td>
<td>0.62</td>
<td>5.70</td>
<td>0.56</td>
<td>13.91</td>
<td>0.99</td>
<td>34.15</td>
</tr>
<tr>
<td>10:15 AM</td>
<td>10:30 AM</td>
<td>0.94</td>
<td>4.57</td>
<td>1.34</td>
<td>0.69</td>
<td>5.27</td>
<td>1.30</td>
<td>5.27</td>
<td>1.30</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>10:45 AM</td>
<td>9.01</td>
<td>2.34</td>
<td>1.27</td>
<td>0.51</td>
<td>2.19</td>
<td>1.21</td>
<td>2.19</td>
<td>1.21</td>
</tr>
<tr>
<td>10:45 AM</td>
<td>11:00 AM</td>
<td>1.49</td>
<td>16.17</td>
<td>1.01</td>
<td>0.91</td>
<td>2.61</td>
<td>1.31</td>
<td>2.61</td>
<td>1.31</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>11:15 AM</td>
<td>13.48</td>
<td>0.98</td>
<td>1.94</td>
<td>1.14</td>
<td>0.64</td>
<td>5.19</td>
<td>1.40</td>
<td>5.19</td>
</tr>
<tr>
<td>11:15 AM</td>
<td>11:30 AM</td>
<td>3.93</td>
<td>0.80</td>
<td>1.50</td>
<td>1.25</td>
<td>0.97</td>
<td>1.85</td>
<td>1.25</td>
<td>1.85</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>11:45 AM</td>
<td>3.80</td>
<td>1.11</td>
<td>0.68</td>
<td>2.97</td>
<td>1.15</td>
<td>2.97</td>
<td>1.15</td>
<td>2.97</td>
</tr>
<tr>
<td>11:45 AM</td>
<td>12:00 PM</td>
<td>3.40</td>
<td>7.74</td>
<td>0.92</td>
<td>5.16</td>
<td>1.26</td>
<td>0.78</td>
<td>3.53</td>
<td>1.51</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>12:15 PM</td>
<td>3.38</td>
<td>9.75</td>
<td>1.07</td>
<td>3.93</td>
<td>1.56</td>
<td>0.63</td>
<td>10.15</td>
<td>1.28</td>
</tr>
</tbody>
</table>

**a) AM Period**

<table>
<thead>
<tr>
<th>Segment</th>
<th>1:2</th>
<th>2:3</th>
<th>3:4</th>
<th>4:5</th>
<th>5:6</th>
<th>6:7</th>
<th>7:8</th>
<th>8:9</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin Time</td>
<td>End Time</td>
<td>21.18</td>
<td>2.66</td>
<td>4.67</td>
<td>3.15</td>
<td>1.17</td>
<td>4.96</td>
<td>1.23</td>
<td>4.96</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>5:15 PM</td>
<td>3.48</td>
<td>15.59</td>
<td>2.15</td>
<td>4.61</td>
<td>3.23</td>
<td>1.22</td>
<td>8.66</td>
<td>1.27</td>
</tr>
<tr>
<td>5:15 PM</td>
<td>5:30 PM</td>
<td>3.54</td>
<td>14.31</td>
<td>1.52</td>
<td>5.43</td>
<td>2.85</td>
<td>1.20</td>
<td>6.40</td>
<td>1.25</td>
</tr>
<tr>
<td>5:30 PM</td>
<td>5:45 PM</td>
<td>14.54</td>
<td>1.55</td>
<td>5.54</td>
<td>1.63</td>
<td>1.02</td>
<td>15.86</td>
<td>1.24</td>
<td>15.86</td>
</tr>
<tr>
<td>5:45 PM</td>
<td>6:00 PM</td>
<td>10.90</td>
<td>1.42</td>
<td>4.46</td>
<td>1.72</td>
<td>1.18</td>
<td>16.54</td>
<td>1.27</td>
<td>16.54</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>6:15 PM</td>
<td>3.63</td>
<td>8.54</td>
<td>1.55</td>
<td>2.97</td>
<td>1.30</td>
<td>0.91</td>
<td>20.08</td>
<td>1.22</td>
</tr>
<tr>
<td>6:15 PM</td>
<td>6:30 PM</td>
<td>3.17</td>
<td>7.39</td>
<td>0.98</td>
<td>2.21</td>
<td>1.18</td>
<td>1.01</td>
<td>27.88</td>
<td>1.29</td>
</tr>
<tr>
<td>6:30 PM</td>
<td>6:45 PM</td>
<td>4.05</td>
<td>8.57</td>
<td>0.94</td>
<td>1.74</td>
<td>1.14</td>
<td>0.88</td>
<td>16.67</td>
<td>1.21</td>
</tr>
<tr>
<td>6:45 PM</td>
<td>7:00 PM</td>
<td>8.32</td>
<td>1.23</td>
<td>2.23</td>
<td>1.19</td>
<td>0.63</td>
<td>2.25</td>
<td>1.18</td>
<td>17.02</td>
</tr>
<tr>
<td>7:00 PM</td>
<td>7:15 PM</td>
<td>3.34</td>
<td>9.47</td>
<td>1.18</td>
<td>1.93</td>
<td>1.09</td>
<td>0.58</td>
<td>14.70</td>
<td>1.20</td>
</tr>
<tr>
<td>7:15 PM</td>
<td>7:30 PM</td>
<td>3.17</td>
<td>8.05</td>
<td>1.02</td>
<td>2.36</td>
<td>1.06</td>
<td>0.50</td>
<td>52.79</td>
<td>1.28</td>
</tr>
<tr>
<td>7:30 PM</td>
<td>7:45 PM</td>
<td>9.04</td>
<td>0.96</td>
<td>1.68</td>
<td>1.09</td>
<td>0.71</td>
<td>21.26</td>
<td>1.24</td>
<td>70.23</td>
</tr>
<tr>
<td>7:45 PM</td>
<td>8:00 PM</td>
<td>1.07</td>
<td>0.64</td>
<td>2.10</td>
<td>1.29</td>
<td>1.29</td>
<td>1.29</td>
<td>1.29</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Table 1: Route 1 Travel Times, January 15th, 2011 (minutes)

**Route Travel Time Comparison**

While nine separate routes were designated as possible alternatives, only three, routes 1, 3, and 4, those shown in Figure 2, had sufficient data for analysis. Figure 6 shows a comparison plot for these three routes, with the raw data in Figure 6a.
Figure 6: Complete Route Travel Time Comparison, Saturday, January 15th, 2011
It is challenging to identify any trend in Figure 6a, due to the variance of one segment shared by routes 1 and 3, highlighted in Table 1b. One possible treatment for this variance is to substitute a calculated, assumed travel time value for the segment. While this does not solve the problem of variance, it does allow the analyst to more easily compare the route travel times to other performance measures. From experience, the research staff was aware of a bottleneck upstream of this segment, resulting in traffic that is typically free flowing in this segment. A reasonable travel time of 5 minutes was substituted for the segment 7:8 data. This data is plotted in Figure 6b.

With the estimation method applied to the data, the travel times traces from the vendor software output on all three studied routes are relatively close to each other, though there does appear to somewhat more variance on Route 4. This is probable due to the fact that only one segment, 2:3, which is shared by all three routes, has the possibility to markedly increase route travel time. Other segments that are unique to individual routes, while they may experience congestion and contribute to an increased travel time, cannot provide an increase on the order of segment 2:3, helping to keep the reported travel times for the three routes relatively close to each other.

Case Study Wrap-up

With the data shown in Figure 6, along with route comparisons of other analysis days, if enough MAC address matches are available to generate output from the software, the travel time data provided by the vendor software in this study appears to be reasonable, given the expected peak period travel times along the studied routes. While the times provided may not be on the order of the anecdotal evidence collected during the winter holidays of 2010, the comparison of the output from the post-processing to independently developed data lends credence to the data supplied by the vendor software. The caveat, however, is that on the lesser travelled segments of the routes studied in this work there is rarely enough data to provide a route travel time. Given this, along with the fact that the reported travel times along the three studied routes in this work are relatively similar to each other, the research staff did not recommend to proceed to future phases on this project (providing real time travel time feedback to users) using only Bluetooth MAC address tracking technology for generation of travel time data.

Lessons Learned and Conclusions

With hindsight, the natural question to ask is ‘What could we have done better?’ To start, the research staff could have asked better questions of potential vendors at the onset of the project. Data insufficiency was a problem with this work, and the algorithms within the vendor software were not able to consistently estimate travel times with the low amount of data. For one embarking on a similar project, the authors suggest asking the following questions of a vendor.

1. What data has been used for your algorithm development?
2. Was it primarily from a freeway, rural highway, urban arterial, or other?
3. What was the AADT on that facility?
4. What is the expected penetration rate?
5. In the software, what specific values are user adjustable?
Answers to these questions might provide insight as to whether or not a specific vendor’s solution will be applicable to the problem at hand. Also, it might be helpful to ask for a software walkthrough from the vendor, to ensure the software has the features required for the problem to be addressed.

Additionally, performing a comparison of the AADT of the facilities to be studied with the expected sensor penetration rate may provide insight as to whether or not enough data will be present to conduct an analysis. Such a comparison would also help in determining sensor placement, as locations that may be questionable could be avoided and, as needed, data collected by other means, e.g., probe vehicle.

Conclusions

The literature review and case study presented in this document have raised important issues regarding the development of accurate travel times with Bluetooth MAC address tracking. A lack of data hampered results in the case study, in which travel times were not available throughout much of the study area due to data insufficiency. In addition, extreme variance was observed on one segment due to the inability of the supplied software to remove outliers from the data, an issue which was likely exacerbated by insufficient data. While the generation of travel times with Bluetooth MAC address tracking has been shown to be a reliable surrogate for alternate measures of determining travel time, it is not a panacea for all situations. Serious consideration should be paid to the issues raised in the literature review and case study prior to embarking on a Bluetooth deployment. On the research side, the authors recommend future work in the area of low volume algorithms as well as non-vehicular modes of travel. While it is understood that there is no substitute for actual data points, additional research focusing on how to handle small amounts of data may have helped provide better results in this case study.


5 Quayle, S.M., Koonce, P., DePencier, D., and Bullock, D.M. “Arterial Performance Measures with Media Access Control Readers.” Transportation Research Record: Journal of the


12 BluSTATS Operations Manual, Traffax, INC. February 2012