

Fundamental Characteristics of Wi-Fi and Wireless Local Area Network Re-identification for Transportation

Noah J. Goodall, Ph.D., P.E.

Virginia Transportation Research Council, 530 Edgemont Road, Charlottesville, VA USA

noah.goodall@vdot.virginia.gov

*This paper is a postprint of a paper submitted to and accepted for publication in IET Intelligent Transport Systems and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at the IET Digital Library:
<http://dx.doi.org/10.1049/iet-its.2016.0087>.*

Abstract: Many transportation agencies use re-identification technologies to identify vehicles at multiple points along the roadway as a way to measure travel times and congestion. Examples of these technologies include license plate readers, toll tag transponders, and media access control (MAC) address scanners for Bluetooth devices. Recent advancements have allowed for the detection of unique MAC addresses from Wi-Fi and wireless local area network (WLAN) enabled devices. This paper represents one of the first attempts to measure the fundamental characteristics of Wi-Fi re-identification technology as it applies to transportation data collection. Wi-Fi sampling rates, re-identification rates, range, transmission success rates, and probability of discovery of sensors and mobile devices were measured, and a model of probability of detection is presented. Field tests found that mobile phones routinely experienced significant time gaps between Wi-Fi transmissions. The study recommends that Wi-Fi sensors be deployed at low-volume, low-speed roadways, with sensors positioned near intersections where vehicles are likely to slow or stop. Due to Wi-Fi's relatively low probability of discovery, the technology may produce poor results in applications that require re-identifying vehicles over multiple consecutive sensors.

1. Introduction

The monitoring of roadway performance is of great importance to transportation agencies. Travel time data are often used for congestion and reliability reporting, real-time display of expected travel times on dynamic message signs, and signal timing studies. One of the most widely used techniques to collect travel time data is vehicle re-identification, where a single vehicle is identified at two separate locations along a roadway. The time difference between those two observations is assumed to be the vehicle's travel time. Several technologies have been used by transportation agencies to re-identify vehicles, including license plate readers, toll tags, and scanning of unique media access control (MAC) addresses from Bluetooth-discoverable mobile devices.

Bluetooth scanning technologies have become widely used since their introduction in the early 2000s. Many transportation agencies rely on Bluetooth sensing as ground truth to validate private sector probe-based travel time data [1]. While Bluetooth sensing has been successful at measuring travel times on

high volume corridors such as urban freeways, at times it has produced poor results on low-volume corridors or during off-peak hours, due to its relatively low sampling rate [2].

In order to increase sample size during low-volume periods, some agencies have begun investigating Wi-Fi sensing technologies. These sensors detect the unique MAC addresses associated with mobile phone, tablet, laptop computer, or any other mobile devices with the ability to transmit a Wi-Fi signal.

Researchers have attempted to use Wi-Fi sensing in various applications. Much of the work has focused on pedestrian characteristics such as movement tracking [3, 4], density [5], flow [6], travel time [7], and demand [8].

Research on re-identification of Wi-Fi devices in vehicles often compares Wi-Fi's performance with the more widely-studied Bluetooth sensors. Several studies have found that Wi-Fi generates more detections at a single location than Bluetooth. Luber et al. [9] studied penetration rates on a German freeway in 2011 found approximately 6.5 unique Bluetooth MAC addresses and 1 Wi-Fi MAC address for every 100 vehicles. Abbot-Jard et al. [10], which found that Bluetooth generated almost twice as many re-identifications as Wi-Fi, but did not measure vehicles volumes to obtain percentages. The authors note, however, that the low counts for Wi-Fi may be due to their Wi-Fi sensor being limited to a maximum of five MAC addresses sensed per scanning cycle [10]. Other studies find that Wi-Fi sensors discover more MAC addresses than Bluetooth sensors: on a lower-speed arterial facility, Lesani et al. [11] detected 16 Bluetooth MAC addresses and 27 Wi-Fi MAC addresses for every 100 vehicles, while Puckett and Vickich [12] found that Wi-Fi sensors detected sixteen times more MAC addresses at single location, but re-identified only twice as many MAC addresses over two consecutive sensors [12]. A study of integrated sensing at Arizona-Mexico ports of entry found that Wi-Fi produced 4-5 times more samples than Bluetooth [13]. Wi-Fi sampling rates varied wildly, typically at 20% but reaching nearly 100% at certain ports and certain times of day [13]. While these studies compared sampling and re-identification rates of Wi-Fi and Bluetooth, they did not investigate device-level communication characteristics such as positioning or transmission rate.

Other studies have investigated Wi-Fi for vehicle monitoring without comparing to Bluetooth. Musa and Eriksson [14] developed and tested a method to estimate trajectories Wi-Fi mobile devices across a network of closely-spaced (400 meters) sensors, and were able to estimate device position with a mean error of 70 meters. Jackson et al. [7] classified mobile devices by mode (pedestrian, cyclist, vehicle) based on estimated speeds over six consecutive Wi-Fi sensors.

The use of re-identification technologies to measure travel times has been covered extensively in the literature, and will not be discussed in this paper. Detailed communications protocols for Bluetooth [15] and Wi-Fi [14] can also be found in the literature and the IEEE 802 standard [16]. Privacy concerns, specifically the challenge of linking a Wi-Fi MAC address with a unique individual, have also been the subject of several studies [17, 18].

Research to date has focused on travel time data collection and sample sizes from Wi-Fi, without investigating the causes of these characteristics as they compare to Bluetooth, nor the reasons for the discrepancies among studies. The general characteristics of Wi-Fi sensing in a transportation environment, especially how they differ from similar Bluetooth technologies, have received little attention in the literature.

The objective of this paper is to present a preliminary investigation into the unique characteristics of Wi-Fi re-identification as they relate to transportation. A secondary objective is to compare Wi-Fi and Bluetooth re-identification. This is the first study to directly compare Wi-Fi and Bluetooth sensing while investigating the properties of Wi-Fi detection at the mobile device level. The results of this study will assist researchers and transportation agencies in determining appropriate use cases for Wi-Fi detection, interpreting data from Wi-Fi sensors, and designing effective deployment plans.

The remainder of this paper is organized as follows. Section 2 investigates the characteristics of Wi-Fi re-identification technology from field tests, and describes a model to predict the probability that a mobile device will be discovered by a sensor. Section 3 compares the characteristics of Wi-Fi re-identification with the more commonly-used Bluetooth technologies. Section 4 describes the author's conclusions.

2. Characteristics of Wi-Fi transmissions

An understanding of the basic characteristics and performance of Wi-Fi data collection are crucial to developing an effective deployment plan. This section describes several tests of a Wi-Fi sensor in lab and field environments that investigate the sensor's performance when detecting transmissions from individual mobile devices. For some tests, only a single phone was used, and readers should use caution when extrapolating findings to a broader population. Readers should also use caution when applying these findings in the future, as mobile technology continues to evolve.

2.1. *Transmission range*

In field tests, a car equipped with a global positioning system (GPS) device was driven slowly around several parking lots within sight of a Wi-Fi sensor. A mobile phone (iPhone 4s running iOS 7) was positioned in the car, while the Wi-Fi sensor remained stationary. In this test, as in all others, the Wi-Fi sensor used a 2.4 GHz antenna with a gain of 3 dBi, impedance of 50 Ohm, and return loss of -20 dB (minimum). The antenna was mounted on a small plywood box, with the receiver resting on top of the box. For the test, the antenna was positioned at a height of approximately 1.5 meters.

To replicate the typical placement of a mobile device in a vehicle, the phone was tested in three positions: held up to the car window as if in use for a phone call, set below the window near in the driver's seat to simulate being in a pocket, and in the car's glove compartment. Only one Wi-Fi sensor was used, at the base station, so signal transmission interval was not measured. The results of the field test are presented in Figure 1. The phone's MAC address was detected at 340 meters, the maximum line of sight tested. Signal strength (as measured by the Wi-Fi sensor) as a function of distance (as measured using GPS) were similar for all three phone positions, suggesting that range and transmission interval are unaffected by the position of the mobile device inside the vehicle.

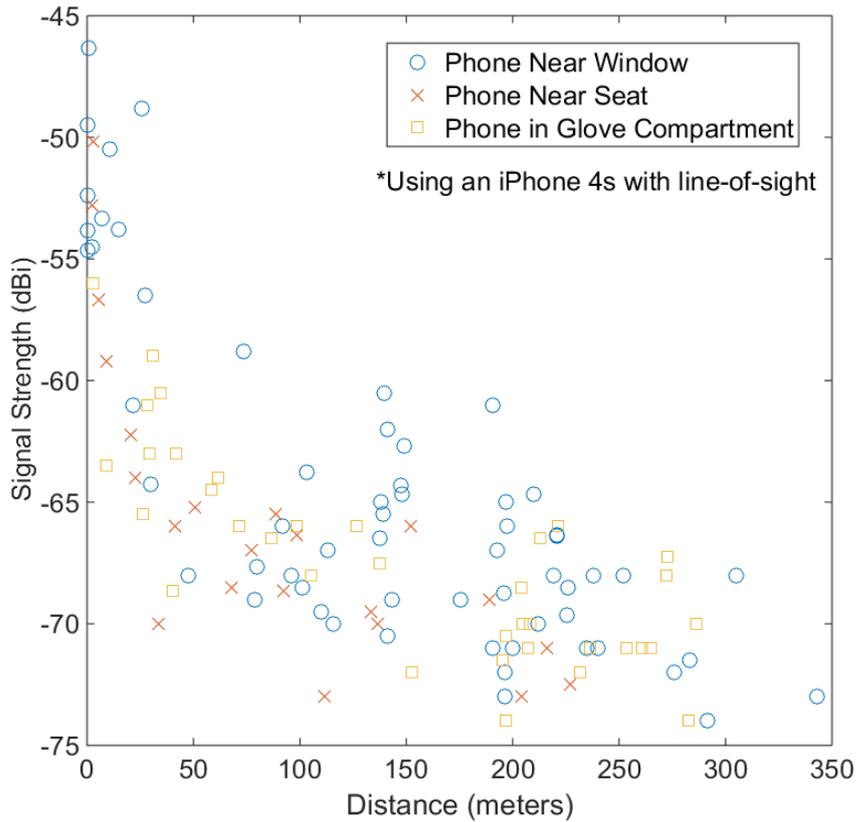


Fig. 1. Observed strength and range of mobile phone Wi-Fi signal

2.2. Transmission interval

Unlike Bluetooth signals, which are transmitted frequently and where sensing is limited by the sensor’s scanning cycle—recommended as once every 10.24 seconds [19]—Wi-Fi signals are comparatively infrequent, often experiencing long gaps between transmissions.

The performance of Wi-Fi detection technology was tested to obtain a baseline rate of MAC address reception. Although the mobile device’s actual transmission interval could not be measured directly, a baseline reception interval was established by positioning a Wi-Fi sensor in the immediate vicinity of the mobile device. Throughout this paper, this baseline reception interval is referred to as the mobile device’s *transmission interval*.

Transmissions arrived in bursts—typically four or five within two seconds—and with gaps of various lengths between bursts. For the purposes of travel time measurements, a burst can be treated as a single transmission, as the benefits of receiving two transmissions within a single burst are negligible.

Based on tests with mobile phones and a nearby Wi-Fi sensor, 57.9% of transmissions occurred within 0.2 seconds of each other, with only 0.3% of transmissions having gaps of between 0.2 and 1 second. From these data, 0.2 seconds was selected as the minimum gap time between two transmissions for them to be considered as unique transmissions. Unless otherwise specified, any analysis of transmission intervals, success rates, and re-identification rates in this paper treats these transmission bursts as individual transmissions.

The time intervals between transmissions were tested on a variety of mobile phones and using several different settings. All phones were placed within one meter of the Wi-Fi sensor antenna in an indoor office environment. In each test, phones were turned off, then turned on and measured over 20 minutes, to simulate a typical commute. Three phones were tested while operating on battery power, three phones were tested while plugged in to wall chargers, and one phone was tested while running both music streaming and navigation apps over the cellular network (with Wi-Fi enabled but without a Wi-Fi connection). A single test was conducted for each scenario. The results of the experiment are shown in Figure 2, where each circle represents the elapsed time since the previous transmission, excluding transmissions that are received within 0.2 seconds of the previous.

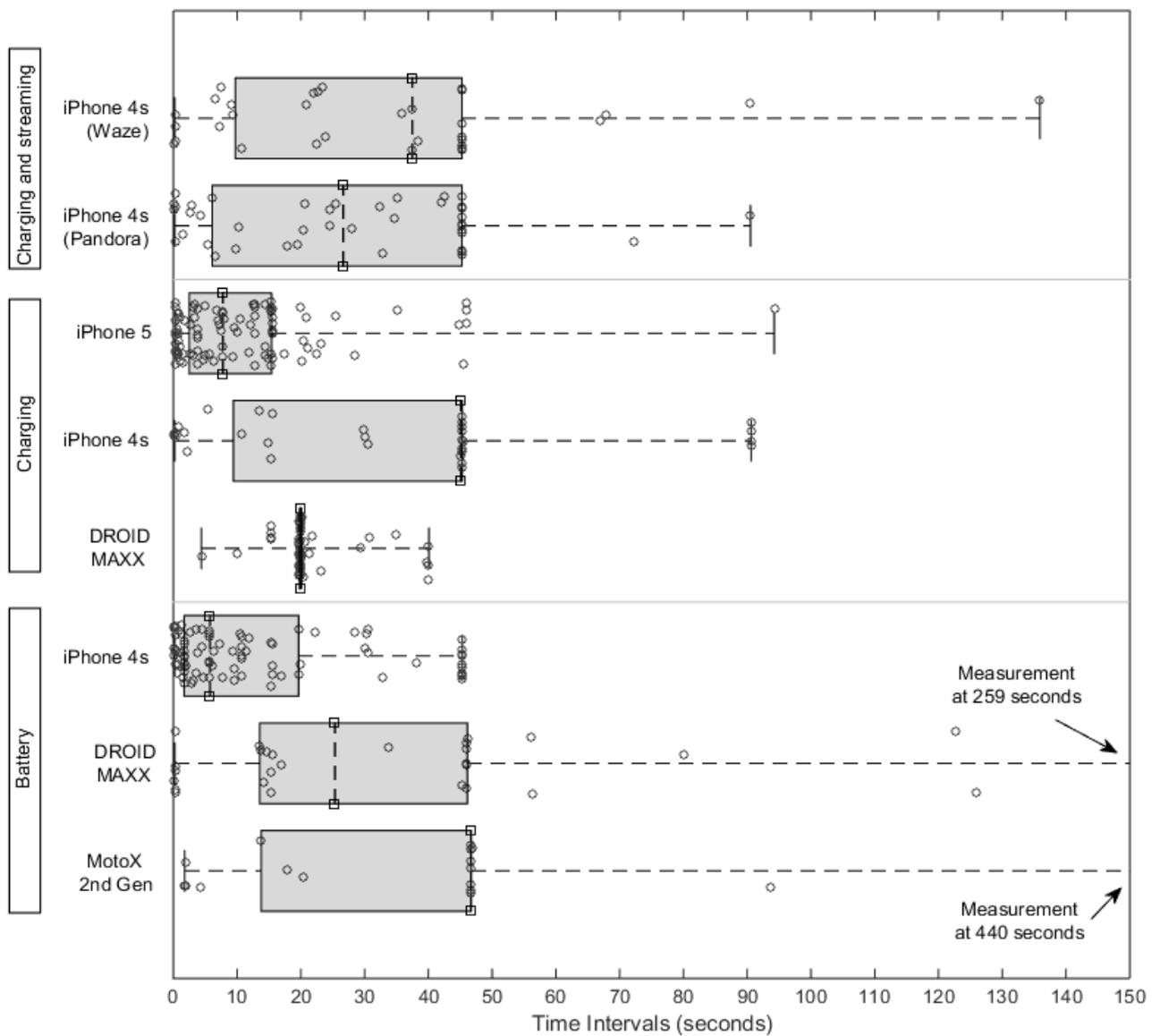


Fig. 2. Baseline Wi-Fi transmission intervals of various mobile phones

Average time intervals between transmissions for an untouched phone after a 10-minute warm-up period (when the phone is first turned on) ranged between 17.7 and 82.0 seconds depending on phone and scenario, with the single longest observed intervals at 7 minutes and 20 seconds. The remaining analysis in this paper assumes a 45 second average gap between transmissions as expected performance.

2.3. Transmission success rate

Although mobile phones were observed transmitting Wi-Fi signals at up to 340 meters with 45 second gaps between transmissions, this does not guarantee that a roadside Wi-Fi sensor will receive signals at this range and rate. To test the detection rate of a mobile device's MAC address transmissions, a second Wi-Fi sensor is needed to act as a baseline to determine when the mobile device is sending signals.

In this experiment, a mobile phone (iPhone 4s running iOS 8.2) was driven along an 8-kilometer section of US-50. The test corridor is a 4-lane, high-volume arterial in Fairfax County, Virginia near Washington, DC with a posted speed limit of 72 km/hr (45 mi/hr). Five intersections along the corridor were equipped with Wi-Fi sensors that detected signals from the mobile phone as it travelled past. A map of the corridor is shown in Figure 3.



Fig. 3. Map of the study area in Fairfax County, Virginia

To determine the phone's distance from a given Wi-Fi sensor, a GPS device collected and logged the position of the test vehicle. These positions were later used to determine the vehicle's distance from the known coordinates of each of the five roadside Wi-Fi sensors. To record a baseline of the phone's transmission interval, a separate Wi-Fi sensor was positioned inside the test vehicle, within one meter of the mobile phone. The timestamps of the GPS device, in-vehicle Wi-Fi sensor, and roadside Wi-Fi sensors were manually synchronized at the beginning of the test to ensure that all data were directly comparable. The test vehicle was driven for 1.5 hours along the corridor, stopping within 20 meters of each Wi-Fi unit for three to five minutes. As with the transmission interval tests, any transmissions within 0.2 seconds of each other were considered a single transmission.

It should be noted that Bluetooth and Wi-Fi sensors were located within 50 meters of each other at each of the five measurement sites on US-50, and that this may have caused interference. The Wi-Fi sensors in this study were found to capture 15% fewer matched pairs on a day with Bluetooth sensors present (Tuesday, July 7, 2015) than a comparative day without Bluetooth sensors (Tuesday, June 30, 2015). The sensor at Cherry Street showed the greatest difference, with a 36% decrease in matched pairs when the Bluetooth sensor was present. When the Cherry Street sensor was excluded from the data, the other sensors experienced only a 3% decrease on days when co-located with Bluetooth sensors. Because it is difficult to compare Bluetooth and Wi-Fi without co-locating sensors, the full data set is included in the analysis. Observed sampling rates should be considered a baseline of Wi-Fi's potential given the possible interference from Bluetooth sensors.

During the field test, the in-car Wi-Fi sensor received the mobile phone's MAC address 378 times while within 200 meters of a roadside Wi-Fi sensor. Of these transmissions, 56 were also recorded by a roadside unit, for a success rate of 14.8%. The signal was more likely to be recorded when sent over shorter distances, with a 21.9% success rate when under 50 meters, compared to 2.8% from between 100 and 200 meters. The results of the transmission success rate test are shown in Table 1.

Table 1 Observed transmission success rate of Wi-Fi mobile device

Distance (<i>m</i>)	Transmission Success Rate	Transmissions Sent
0 - 50	21.9%	196
50 - 100	14.9%	67
100 - 150	2.8%	72
150 - 200	2.3%	43

2.4. Model of probability of discovery

Assuming a known mobile device transmission interval and the probability of a roadside Wi-Fi sensor receiving a transmission, one can develop a general model for predicting the probability of discovering a mobile device. A model of the probability of discovery is proposed here, which may prove useful to those deploying Wi-Fi sensors and wishing to track vehicles over multiple sensors.

This model begins with the assumption that a subject Wi-Fi device is within range of a roadside sensor, and is occasionally transmitting packets at some frequency. The probability of the Wi-Fi device being detected by the roadside sensor, $P(x)$, requires three pieces of information: the length of time (T) that

the vehicle (x) is within range of the roadside sensor, the mobile device's transmission frequency (ω), and the transmission success rate (ρ). The probability of a roadside sensor not receiving a given transmission ($1 - \rho$) is multiplied by itself for every expected transmission sent ($T\omega$). This value is then subtracted from 1 to find the probability of receiving a transmission and discovering the device. The probability of discovery can be calculated using Equation 1.

$$P(x) = 1 - (1 - \rho)^{T\omega} \quad (1)$$

Because a roadside sensor is more likely to discover a mobile device the longer it is in range, roadside sensors may be located at traffic signals where vehicles experience some delay. When the Wi-Fi sensor is located at an intersection, the probability can be defined as having two components. The first component is the probability of discovery when the vehicle is moving at free flow speeds, and the second is when the vehicle is delayed by traffic or the traffic signal. For simplicity, assume that all delay occurs within 50 meters of the traffic signal. Therefore, the two components have different times within range (T) and transmission success rates (ρ), as success rate is higher when closer to the sensor. The time within range for the free flow component (T_{ff}) can be described as the travel distance within range ($2r$) divided by speed (v). This relationship is shown in Equation 2.

$$T_{ff} = \frac{2r}{v} \quad (2)$$

In this case, free flow travel time should be multiplied by the average transmission success rate over the sensor's range ($\bar{\rho}$). The delay component can be calculated similarly to Equation 1, but instead using the delay time (T_{delay}) in place of travel time, and the transmission success rate within 50 meters of the sensor ($\rho_{<50}$) instead of the average transmission success rate. The full model, when the roadside sensor is located at the traffic signal, is described in Equation 3.

$$P(x) = 1 - \left[(1 - \bar{\rho})^{\frac{2r\omega}{v}} \times (1 - \rho_{<50})^{T_{delay}\omega} \right] \quad (3)$$

Alternatively, in scenarios where the Wi-Fi sensor is not co-located at a signalized intersection, and where vehicle speeds through the sensor's coverage area are expected to be relatively uniform, a simplified form of the model can be used. Equation 4 describes the simplified model, where vehicle speed (v) represents the average speed instead of free flow speed. This model is similar to Equation 3, but excludes the intersection delay component.

$$P(x) = 1 - (1 - \bar{\rho})^{\frac{2r\omega}{v}} \quad (4)$$

Many of the terms in the probability of discovery model can be determined from field measurements. Based on observations from this study, the values shown in Table 2 were used to calculate probabilities of discovery using Wi-Fi sensors. It should be noted that the values in Table 2 are specific to this study's test conditions and tested devices, and may not be representative of the conditions or Wi-Fi mobile device population of a different roadway.

Table 2 Probability of mobile device discovery model: terms and observed values

Symbol	Description	Observed Values
r	Sensor's effective range	200 meters
v	Free flow speed (average speed in simplified model)	20.1 m/s (45 mi/hr)
\bar{p}	Average transmission success rate over sensor's full range	14.8%
$p_{<50}$	Transmission success rate within 50 meters of the signal/sensor	21.9%
T_{delay}	Intersection movement delay	Variable
ω	Transmission frequency	0.022 Hz

Table 3 lists expected probabilities of discovering a mobile phone that is transmitting Wi-Fi signals, as calculated using Equation 3 for various free flow speeds and intersection movement delays. As stated earlier, the probabilities listed in Table 3 are dependent upon the Wi-Fi sensor being located near the intersection, as well as the accuracy of the values in Table 2. From the model, the probability of a mobile device being discovered on at 110 km/hr with no delay may be as low as 4.5%, and increasing to 61.0% at 40 km/hr and 150 seconds of delay.

Table 3 Estimated probabilities of discovery based on the model predictions

Free Flow Speed (km/hr)	Movement Delay (s)									
	0s	10s	20s	30s	40s	50s	60s	90s	120s	150s
40	11.9%	16.6%	21.0%	25.2%	29.1%	32.9%	36.4%	46.0%	54.1%	61.0%
50	9.7%	14.4%	19.0%	23.3%	27.3%	31.2%	34.8%	44.6%	53.0%	60.0%
60	8.1%	13.0%	17.6%	21.9%	26.1%	30.0%	33.7%	43.7%	52.2%	59.4%
70	7.0%	11.9%	16.6%	21.0%	25.2%	29.1%	32.9%	43.0%	51.6%	58.9%
80	6.1%	11.1%	15.8%	20.3%	24.5%	28.5%	32.3%	42.5%	51.1%	58.5%
90	5.5%	10.5%	15.2%	19.7%	24.0%	28.0%	31.8%	42.1%	50.8%	58.2%
100	4.9%	10.0%	14.7%	19.3%	23.5%	27.6%	31.4%	41.7%	50.5%	58.0%
110	4.5%	9.6%	14.3%	18.9%	23.2%	27.2%	31.1%	41.5%	50.3%	57.8%

Musa and Eriksson [14] developed a similar model for the probabilities of detecting a mobile device at different distances from a sensor, which they refer to as emissions probabilities. The motivation for their effort was to re-construct the trajectories of mobile devices moving through a street network with irregularly spaced sensors; if a mobile device was detected at the first and third sensor, but not the second, it is useful to know the probability of non-detection as it suggests the probability that a mobile device travelled along an alternate route. Their model differs from the one presented here in several ways: it incorporates the signal strength of the received transmission, it does not assume that a mobile device passes within range of a sensor for any length of time, and requires the experimental measurements of transmission probabilities from devices with known transmission rates. Their approach is favourable for trajectory reconstruction over a network, whereas the approach described in this paper is more straightforward while remaining suitable for a corridor study.

2.5. Empirical probability of discovery

In order to validate the model's estimated probabilities of discovery, data from the existing roadside sensors were evaluated to determine the percentage of Wi-Fi equipped devices travelling past the sensors that were detected. This can be challenging to measure, as it requires knowledge of the total number of Wi-Fi enabled mobile devices passing by a sensor. Although one could compare the number of Wi-Fi devices detected at two adjacent sensors, there is no guarantee that the device detected at the first sensor has also travelled to the second. There are several reasons for this: the device may not be moving, e.g. a

laptop in a nearby office building; the device may be travelling away from the second sensor, either on the main road or on a side street; or the device passes the second sensor but is never detected because it does not transmit its MAC address or its signal is not received.

In order to identify a set of mobile device MAC addresses that have travelled past a roadside sensor but may have gone undetected, the records from sensors at opposite ends of the 8-kilometer corridor were searched to find matching MAC addresses within a reasonable travel time window. Because there is no obvious alternate route through the corridor using parallel roads, it was assumed that mobile devices seen at both ends of the corridor were not only travelling in vehicles, but also passed each of the three roadside Wi-Fi sensors in the time between their first and last recorded detections at the terminus sensors.

The anonymized Wi-Fi MAC addresses of mobile devices were collected by five roadside sensors over eight consecutive days, including both weekday and weekends. Matches were filtered for data quality by comparing each match against the 30 nearest matches, both past and future. The difference between the 75th and 25th percentile of the nearest 30 matches was defined as the record’s inter-quartile range travel time. Records with travel times less than the 25th percentile minus twice the inter-quartile range were excluded. Similarly, records with travel times greater than the 75th percentile plus twice the inter-quartile range were excluded [20]. In our data, 6-18% of matches were removed per link based on these criteria.

Of the 303,190 unique MAC addresses recorded at Nutley Street (the farthest west roadside sensor) and the 548,858 recorded at Cherry (the farthest east), 2,619 addresses were identified at both sensors (0.86% and 0.48% respectively). Of the 2,619 devices believed to have travelled through the corridor, the three remaining sensors discovered between 36% and 63% of the original devices. For comparison, a similar technique was used with Bluetooth sensors co-located with the Wi-Fi sensors. Bluetooth units detected a similar number of through-vehicles (2,528) from a lower number of total records (13.4% of 18,918 records at Nutley Street, and 13.4% of 18,891 at Cherry Street). However, the three Bluetooth units within the corridor re-identified between 75% to 88% of these through vehicles. The results of the analysis are shown in Table 4.

Table 4 Discovery rate of through vehicles for Bluetooth vs. Wi-Fi

Cross Street	Wi-Fi		Bluetooth	
	<i>Discovered</i>	% of <i>n</i>	<i>Discovered</i>	% of <i>n</i>
At both ends	<i>n</i> = 2619	100%	<i>n</i> = 2528	100%
Prosperity Ave.	1660	63%	2049	81%
Allen St.	1454	56%	2231	88%
Annandale Rd.	950	36%	1898	75%

The speed limit along the section of US-50 where the roadside sensors were placed is 72 km/hr (45 mi/hr). Based on the Wi-Fi reported travel times of the 2,619 through vehicles, each vehicle's delay through the corridor could be determined by subtracting 400 seconds (the free flow travel time) from its observed travel time. Average delay experience by a vehicle at a signal could be calculated by dividing total delay experienced on the trip by the number of traffic signals along the corridor. The probability of discovery for each vehicle could then be estimated using Equation 3. The average total trip delay was nearly 12 minutes, with average delay per signal at 60 seconds.

Using Equation 3 for all vehicles, the average expected probability of discovery via Wi-Fi was 32.7%. This estimate was lower than the observed 36-63% rediscovery rate. This difference may be due to several factors. First, the sample set of MAC addresses may be from devices that are recorded by roadside units more frequently than the general population. By limiting the sample set to vehicles that had already been identified by two roadside sensors, it may have biased the population with more easily-detectable devices. Second, the Wi-Fi transmission frequency and/or the transmission success rate may be lower than that of many devices on the road. For example, alternative devices such as laptop computers and tablets may send out stronger, more frequent signals than the tested mobile phones, which would result in higher re-identification rates. Finally, the delay estimates may be inaccurate, with the sensor locations experiencing greater than their share of the total corridor delay.

A study in Australia found that 19% of all Wi-Fi signals detected at one of two sensors also appeared at the other sensor, compared to 81% for Bluetooth devices [10]. The results presented in this paper found similar Bluetooth discovery rates compared to [10], but much higher Wi-Fi discovery rates. This discrepancy suggests that the sample of Wi-Fi devices used in this paper may have been biased towards more easily-detectable devices.

3. Wi-Fi and Bluetooth comparisons

Wi-Fi sensors are often marketed as a solution to the problem of low penetration rates experienced with Bluetooth sensors. Studies have shown sampling rates for Bluetooth devices as low as 2% to 3.4% [2]. On low-volume roads and during off-peak periods, this can result in sample sizes of Bluetooth-reported travel times that are too small to be reliable. This section describes differences in the sampling rates for Wi-Fi sensors as compared to Bluetooth sensors at both single sensors and matches across multiple sensors.

3.1. Sampling rate at a single sensor

The sampling rate, defined here as the number of unique MAC addresses observed per vehicle, is a useful metric for understanding the number of mobile devices that may provide travel times along a corridor. Note that the sampling rate should not be interpreted as the percentage of vehicles that are detected, as a single vehicle may contain two or more mobile devices.

Vehicles were manually counted at two sites and compared with Wi-Fi MAC addresses collected over the same time periods. Data were collected at two sites. The first site was a four-lane arterial (US-29/Seminole Trail) at the intersection of Town Center Drive outside of Charlottesville, Virginia. This site was chosen as it experiences relatively high speeds of 72 km/hr (45 mi/hr) and high volumes. The second site was an isolated four-way stop intersection of two-lane roads (Old Lynchburg Road and Red Hill Road in North Garden, Virginia). Data were collected for two hours at both sites. MAC addresses with greater than 5 minutes between their first and last records were assumed to be non-mobile devices, and were excluded from the analysis. The sampling rates were 44 unique MAC addresses per 100 vehicles at the low-volume site, and 29 at the high volume site. The results suggest a higher proportion of vehicles may be identified at a single sensor from Wi-Fi data than from Bluetooth. This is in contrast to a previous study which found higher rates with Bluetooth devices [10].

3.2. Re-identification rates

A travel time measurement requires the successful matching of a MAC addresses observed at two separate sensors. The study corridor on US-50 in Fairfax County, Virginia was equipped with both portable Bluetooth sensors located at the same intersections as the Wi-Fi sensors. Table 5 compares the total number of matched pairs, each representing a valid travel time record, from Wi-Fi and Bluetooth detection. Wi-Fi sensors recorded almost twice as many matched pairs of MAC addresses than Bluetooth sensors. This suggests that Wi-Fi sensors provide more data points than Bluetooth sensors, although the quality of the travel time data was not investigated as part of this study.

Table 5 Counts of matched pairs of Wi-Fi and Bluetooth devices on US-50

Direction	Wi-Fi Matched Pairs	Bluetooth Matched Pairs	Wi-Fi : BT Ratio
Eastbound	38,273	20,518	1.87 : 1
Westbound	36,446	18,666	1.95 : 1
Total	74,719	39,184	1.91 : 1

3.3. Re-identification across multiple consecutive sensors

Bluetooth mobile device MAC addresses are recorded by sensors with a high degree of reliability, with up to 99% probability of discovery when within range of the sensor for six seconds [21]. In contrast, Wi-Fi devices may remain ‘silent’ for up to a minute or more. This is a potential cause of the low probability of discovery for Wi-Fi devices when compared to Bluetooth devices as shown in Table 4.

As a result, Wi-Fi devices have a much lower probability of being detected at multiple consecutive sensors. Table 6 shows the percentage of devices that were continuously detected starting at a terminus of the study corridor on US-50. For Wi-Fi devices, only 0.13% of devices detected at the first sensor were later detected at all of the following four sensors, compared to 6.67-6.99% for Bluetooth devices. Overall numbers of multiple-detected vehicles were lower for Wi-Fi as well, with 2,583 vehicles detected at all five sensors using Bluetooth compared to 1,043 for Wi-Fi. This finding suggests that Bluetooth sensors may be preferable for studies which require the tracking of individual vehicles over multiple sensors.

Table 6 Re-identification of mobile devices over multiple consecutive sensors

Cross Street	Distance (km)	Wi-Fi		Bluetooth	
		Re-identified MAC IDs	% Re-identified from initial sensor	Re-identified MAC IDs	% Re-identified from initial sensor
<i>Eastbound</i>					
Nutley	0	303,190	100%	18,918	100%
Prosperity	1.8	11,619	3.83%	4,268	22.56%
Allen	5.3	2,014	0.66%	1,847	9.76%
Annandale	7.6	676	0.22%	1,456	7.70%
Cherry	8.0	394	0.13%	1,262	6.67%
<i>Westbound</i>					
Cherry	0	548,858	100%	18,891	100%
Annandale	0.4	10,808	1.97%	5,566	29.46%
Allen	2.7	3,162	0.58%	4,188	22.17%
Prosperity	6.2	988	0.18%	1,578	8.35%
Nutley	8.0	649	0.12%	1,321	6.99%

4. Conclusions

This paper presented one of the first in-depth analyses of Wi-Fi MAC address sensing for use in transportation. Significant differences were found in the transmission intervals of Wi-Fi devices compared to Bluetooth devices, as Wi-Fi MAC addresses were observed less frequently and with longer gaps

between transmissions (often 45 seconds or more). The transmission success rate of Wi-Fi devices was also observed to be much lower than found in previous studies' measurements of Bluetooth devices, with an average success rate of 14.8%. A model of a Wi-Fi device's probability of discovery was developed based on sensor range, transmission frequency, vehicle speed, intersection delay, and transmission success rate. Using values from the field tests, the model was found to underestimate the probability of discovery as measured on the US-50 test corridor (32.7% predicted, compared to 36-63% observed). Several possible explanations were discussed. Additional data from the US-50 test corridor revealed that Wi-Fi sensors detected and re-identified more MAC addresses than Bluetooth sensors, but re-detect fewer vehicles over three or more consecutive sensors.

The findings suggest that Wi-Fi sensors may be useful for data collection, especially on low-speed and/or low-volume roads where Bluetooth sensors do not produce adequate sample sizes. Wi-Fi sensors may have difficulty in studies that require re-identifying individual vehicles over multiple consecutive sensors. Because Wi-Fi transmissions are sent infrequently, Wi-Fi sensors should be positioned where vehicles are slowed or stopped to allow adequate time to record a transmission.

5. Acknowledgements

This work was supported by the Virginia Department of Transportation. The author thanks Lewis Woodson for assistance with data collection, Michael Fontaine for insightful commentary, and Acyclica for supplying equipment.

6. References

- 1 Haghani, A., Hamed, M., Sadabadi, K., Young, S., Tarnoff, P.: 'Data collection of freeway travel time ground truth with Bluetooth sensors', *Transp. Res. Rec. J. Transp. Res. Board*, 2010, **2160**, pp. 60–68.
- 2 Sharifi, E., Hamed, M., Haghani, A., Sadrsadat, H.: 'Analysis of vehicle detection rate for Bluetooth traffic sensors: a case study in Maryland and Delaware', *Proc. 18th World Cong. Intell. Transp. Systems*, 2011
- 3 Abedi, N., Bhaskar, A., Chung, E.: 'Bluetooth and Wi-Fi MAC address based crowd data collection and monitoring: Benefits, Challenges and Enhancement', *Proc. Australasian Trans. Res. Forum*, Brisbane, Australia, Oct. 2013
- 4 Chon, Y., Kim, S., Lee, S., Kim, D., Kim, Y., Cha, H.: 'Sensing WiFi packets in the air: practicality and implications in urban mobility monitoring', *Proc. 2014 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2014, pp. 189–200

- 5 Schauer, L., Werner, M., Marcus, P.: 'Estimating crowd densities and pedestrian flows using Wi-fi and Bluetooth', Proc. 11th Int. Conf. Mobile and Ubiquitous Systems: Computing, Networking and Services, Brussels, Belgium, 2014, pp. 171–177
- 6 Fukuzaki, Y., Mochizuki, M., Murao, K., Nishio, N.: 'A pedestrian flow analysis system using Wi-Fi packet sensors to a real environment', Proc. 2014 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing: Adjunct Publication, New York, NY, USA, 2014, pp. 721–730
- 7 Jackson, S., Lesani, A., Miranda-Moreno, L.F.: 'Towards a WIFI system for traffic monitoring in different transportation facilities', Proc. 93rd Annual Meeting of the Transp. Res. Board, Washington, DC, USA, Jan. 2014
- 8 Danalet, A., Farooq, B., Bierlaire, M.: 'A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures', Transp. Res. Part C Emerg. Technol., 2014, **44**, pp. 146–170
- 9 Luber, A., Junghans, M., Bauer, S., Schulz, J.: 'On measuring traffic with Wi-Fi and Bluetooth', Proc. 18th World C. on Intell. Transp. Systems, Orlando, FL, USA, 2011
- 10 Abbott-Jard, M., Shah, H., Bhaskar, A.: 'Empirical evaluation of Bluetooth and Wifi scanning for road transport', Proc. Australasian Trans. Res. Forum, Brisbane, Australia, Oct. 2013
- 11 Lesani, A., Romancyshyn, T., Miranda-Moreno, L.F.: 'Arterial Traffic Monitoring Using Integrated Wi-Fi-Bluetooth System', in Transportation Research Board 95th Annual Meeting, (2016)
- 12 Puckett, D., Vickich, M.: 'Traffic data collection using Wi-Fi MAC address matching', N. Amer. Travel Monit. Expo. Conf. (NATMEC), Chicago, IL, USA, June, 2014
- 13 Lee Engineering, LLC, 'Analysis of Bluetooth and Wi-Fi Technology to Measure Wait Times of Personal Vehicles at Arizona-Mexico Ports of Entry' (Arizona Department of Transportation, 2015)
- 14 Musa, A.B.M., Eriksson, J.: 'Tracking unmodified smartphones using Wi-fi monitors', Proc. 10th ACM Conf. Embedded Net. Sens. Systems, New York, NY, USA, 2012, pp. 281–294
- 15 Bhaskar, A., Chung, E.: 'Fundamental understanding on the use of Bluetooth scanner as a complementary transport data', Transp. Res. Part C Emerg. Technol., 2013, **37**, pp. 42–72
- 16 IEEE 802.11: 'Local and metropolitan area network standards', 2012
- 17 Cunche, M.: 'I know your MAC address: targeted tracking of individual using Wi-Fi' *J. Comput. Virol. Hacking Tech.*, 2013, **10**, (4), pp. 219–227.
- 18 Cunche, M., Kaafar, M.-A., Boreli, R.: 'Linking wireless devices using information contained in Wi-Fi probe requests', *Pervasive Mob. Comput.*, 2014, **11**, pp. 56–69.
- 19 Bluetooth SIG: 'Specification of the Bluetooth system', 2010
- 20 Traffax Inc., 'BluSTATs Operations Manual' (Traffax Inc., 2014), pp. 12-17
- 21 Peterson, B.S., Baldwin, R.O., Kharoufeh, J.P.: 'Bluetooth inquiry time characterization and selection', *IEEE Trans. Mob. Comput.*, 2006, **5**, (9), pp. 1173–1187