Final Report

January 2018

HIGHWAY TRAVEL TIME ESTIMATION WITH CAPTURED IN-VEHICLE WI-FI MAC ADDRESSES: MECHANISM, CHALLENGES, SOLUTIONS AND APPLICATIONS

Pengfei (Taylor) Li, Ph.D., P.Eng. (co-PI)
Department of Civil and Environmental Engr, Mississippi State University, MI 39762

Pitu B. Mirchandani, Sc.D. (co-PI)
School of Computing, Informatics, and Decision Systems Engineering
Arizona State University, Tempe, AZ 85287

Lei Zhang
Department of Civil and Environmental Engr, Mississippi State University, MI 39762

Li Zhang, Ph.D., P.E., F. ASCE. F. ITE
Department of Civil and Environmental Engr, Mississippi State University, MI 39762

This report is based on the research initiated by co-principal investigators Pengfei (Taylor) Li, Mississippi State University (formerly at Arizona State University) and Pitu B. Mirchandani, and was partially supported by a subcontract from University of Nevada, Reno, to Arizona State University entitled ***Measuring Traffic Performance Using Passive Sensing Technologies on Signalized Arterials (Task C)***, under NDOT Agreement No: UNR-14-60.

SOLARIS Consortium, Tier 1 University Transportation Center
Center for Advanced Transportation Education and Research
Department of Civil and Environmental Engineering
University of Nevada, Reno, NV 89557

DISCLAIMER:

The contents of this report reflect the views of the authors, who are responsible for the facts and accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

EXECUTIVE SUMMARY

Passive sensing technologies have emerged to supplement the traffic performance measurement recently. One successful is the travel time estimation based on in-vehicle Bluetooth Media Access Control (MAC) address capturing and matching at different locations. In this paper, we present a novel approach to measuring the traffic performance on highways based on in-vehicle Wi-Fi MAC address capturing. While this novel approach shares similarities with the Bluetooth-based solution, the Wi-Fi sensors can reduce the measuring errors by up to 90% and collect 300% more valid travel time samples than the Bluetooth sensors. We also design algorithms to screen outliers and estimate dynamic travel times on freeways and arterials respectively. At last, experiments in the field are conducted to evaluate the Wi-Fi sensing areas under different antenna configurations and cross compare the travel time estimation with Wi-Fi sensors and Bluetooth sensors. The findings include the Wi-Fi sensors considerably outperform the Bluetooth sensors in capturing MAC address of the passing vehicles, especially in low-traffic areas.

Key words: Traffic Performance, Travel time estimation, Wi-Fi, Bluetooth, traffic performance.

Highway Travel Time Estimation with Captured In-vehicle Wi-Fi MAC Addresses: Mechanism, Challenges, Solutions and Applications

*Pengfei (Taylor) Li, Ph.D., P.Eng.
Department of Civil and Environmental Engineering,
Mississippi State University, 39762
Tel: 662-325-7518; Email: pengfei.li@cee.msstate.edu

Pitu B. Mirchandani, Ph.D.
Arizona State University,
BYENG 553, 699 S. Mill Ave., Tempe, AZ, 85281
Tel: (520)5911/882; Email: pitu@asu.edu

Lei Zhang
Department of Civil and Environmental Engineering,
Mississippi State University, 39762
Email: lz218@msstate.edu

Li Zhang, Ph.D., P.E., F. ASCE. F. ITE
Department of Civil and Environmental Engineering,
Mississippi State University, 39762
Tel.: 662-325-9838; Email: lzhang@cee.msstate.edu

*: Corresponding author

Word count: 5,183+ (7 Figures) *250+ (2 Table)*250=7,433 words

Abstract:

Passive sensing technologies have emerged to supplement the traffic performance measurement recently. One successful is the travel time estimation based on in-vehicle Bluetooth Media Access Control (MAC) address capturing and matching at different locations. In this paper, we present a novel approach to measuring the traffic performance on highways based on in-vehicle Wi-Fi MAC address capturing. While this novel approach shares similarities with the Bluetooth-based solution, the Wi-Fi sensors can reduce the measuring errors by up to 90% and collect 300% more valid travel time samples than the Bluetooth sensors. We also design algorithms to screen outliers and estimate dynamic travel times on freeways and arterials respectively. At last, experiments in the field are conducted to evaluate the Wi-Fi sensing areas under different antenna configurations and cross compare the travel time estimation with Wi-Fi sensors and Bluetooth sensors. The findings include the Wi-Fi sensors considerably outperform the Bluetooth sensors in capturing MAC address of the passing vehicles, especially in low-traffic areas.

Key words: Traffic Performance, Travel time estimation, Wi-Fi, Bluetooth, traffic performance

1. Introduction

The recent mobility report reveals traffic congestion in the U.S. results in nearly five billion hours of delay to road users. Together with wasted fuel, congestion results in over $100 billion in waste per year (*1*). As such, it is important to continuously measure and then improve traffic mobility and system performance, especially in some hot areas such as congested urban arterials. Traffic performance measurement and traffic state estimation replies on data. In the past five decades, the inductive loops or other equivalent technologies (e.g., video detection) have been the primary data inputs for traffic performance evaluation. Although the inductive loops have been proven reliable and effective, they also have several long-standing drawbacks, such as high installation and maintenance costs. To mitigate the challenges in the inductive loop technology, multiple novel sensing technologies have been adopted in traffic management recently. One successful example is the highway travel time estimation based on the captured unique Bluetooth Media Access Control (MAC) addresses of passing vehicles and matches them at different locations. Also the evaluations of the Bluetooth-based solution have been reported effective in estimating traffic performance on arterials (*2*).

Although the Bluetooth-based travel time estimation solution is successful in most places, several issues have been widely reported due to the physical characteristics of Bluetooth technology. Specifically, the Bluetooth-based travel time estimation solution has large timing errors which is defined as the travel time from a Bluetooth device's actual location to the presumed location where this device is discovered), reportedly up to 11 seconds, and a too small size of valid travel time samples in low-traffic areas. While these issues can be mitigated by enhancing the data processing algorithms, it is also desirable to explore alternative sensing technologies to generate data sources with better quality.

In this paper, the authors explore the potential of Wi-Fi MAC address capturing technology and its application to traffic performance measurement. Capturing Wi-Fi MAC address has increasingly received attentions from academia and companies in different domains. Nowadays, a vast majority of personal mobile electronic devices has Wi-Fi capability to communicate based on the MAC address standards, IEEE 802 (*3*). Like Bluetooth, Wi-Fi MAC address is also composed of six bytes of globally unique hexadecimal numbers which can hardly be associated with personal information and therefore chances of privacy infringements via captured Wi-Fi MAC addresses are low. The major advantages of Wi-Fi MAC capturing over the Bluetooth are that the timing error of Wi-Fi MAC address capturing is very small, reportedly less than 1 second (*4*), and the valid sample size is also much bigger than its Bluetooth counterpart under the same conditions. These new features of Wi-Fi technology will help to overcome those problems in the Bluetooth-based solution as well as can possibly generate additional traffic performance measurements. The rest of this paper is structured as follows:

Literature on applying vehicle ID-matching technologies to travel time estimation and applying Wi-Fi to various applications. Then the wireless communication protocols of Bluetooth and Wi-Fi are described and compared. We will also explain why Wi-Fi sensors have superior performance to the Bluetooth sensors in capturing MAC addresses. Following that, algorithms to screen outlier samples and estimate travel time and queue length along arterials are presented. At last, results of several experiments in the field are reported and analyzed.

2. Comparison of MAC address Capturing between Wi-Fi and Bluetooth Technologies

The Bluetooth technology was initially intended to connect devices, such as printers or keyboards, to computers without cables (*5*). In order to exchange data between devices, there are two main stages for connection establishment. The first stage is called inquiry which allows the master inquirer (the road-side Bluetooth sensor in this context) to discover the possible “slave” devices (in-vehicle Bluetooth devices in this context) within the sensing area; the second stage is called *Page* in which the master informs the “slave” units regarding its identification status and common clock. Then the connection between master and slaves is created for data exchange. Apparently, for the Bluetooth MAC address capturing, only the inquiry stage is needed. Whenever an inquiry begins, the road-side Bluetooth sensors will first generate 32 distinct hopping frequencies according its own clock and General Inquiry Access Code (GIAC) protocol. The 32 distinct hopping frequencies are further divided into two subsets of 16 frequencies. Within each subset, the master takes 8 frequencies to transit inquiry information (Tx) for 625 microseconds per frequency slot and holds the other 8 frequency slots for another 625 microseconds per frequency to listen to any responses from in-vehicle Bluetooth devices (Rx). Therefore, it takes about 10 milliseconds for the roadside Bluetooth device to scan all the 16 frequencies within a subset (16 *625 microseconds). In order to avoid errors, all hopping frequencies within a subset must be inquired or listened 256 times before the road-side Bluetooth sensor hops to the other subset and each subset must be scanned twice to avoid missing any responses from in-vehicle Bluetooth devices. As a result, each Bluetooth discovery stage lasts about 10.24 seconds (10 milliseconds *256*4). This result has been proven in the past through experiments in the field (*6*). It is also worth pointing out that the maximum number of in-vehicle Bluetooth device discovered by the road-side Bluetooth device in each inquiry is seven per the Bluetooth standard, no matter how many discoverable Bluetooth devices within the sensing rage. Excessive discovered Bluetooth MAC addresses will be randomly abandoned.

The Wi-Fi (also known as Wireless Fidelity) standard is defined in standard of IEEE 802.11 for wireless local area network connections. In IEEE 802.11, the Wi-Fi MAC address is defined for access to the wireless physical media (i.e., air). A Wi-Fi device normally needs to exchange its MAC address with the Wi-Fi access point (AP) to get the information of that AP or further seek authentication and association. In most personal portable devices, the Wi-Fi module is configured to continuously transmit its MAC address into the air to search the nearby APs. Note that such process is essentially an unencrypted broadcast and those transmitted data packets can be completely discovered by the third-party Wi-Fi monitor (e.g., the road-side Wi-Fi sensor) and the individual Wi-Fi MAC address can be retrieved as well.

To compare the performance of MAC address capturing between the Wi-Fi sensors and Bluetooth sensors, one should focus on two items: sensors' timing error and valid sample size because high sample bias and low valid sample size are two main obstacles to traffic state estimation improvements with the passive sensing technologies. Table 1 shows the comparison between Wi-Fi-based and Bluetooth-based solutions. Further comparison will be presented in the later sections.

Table 1 Comparison between Bluetooth and Wi-Fi MAC Address Capturing

	Bluetooth	Wi-Fi
Max Timing Error at One Location	±10 seconds	±1 second
Max No. of Discovered MAC Addresses in Each Discovery Round at One Location	7	unlimited
Matched Sample Rate	Low	High
Matched Sample Bias	High	Low

3. Data Screening Process

Both Bluetooth and Wi-Fi road-side sensors will generate exactly the same type of data set (i.e., road segment travel time samples) and, as shown in Figure 1, valid samples will always be contaminated first by the random outliers and sensors' inherent measuring errors during the collection. As such, one main step of data processing is to screen the outliers and reduce the bias for the traffic state estimation later. Both Bluetooth and Wi-Fi sensors in essence fall into the category of automated vehicle identification technology (AVI). While they are more cost-effective than many other similar technologies, they are also more likely to generate measurement errors. The measure errors include not only the aforementioned timing errors but also the positioning errors, defined as the location difference between a vehicles' actual location and the presumed discovered location near the road-side sensors. In addition, vehicles may also slow down, speed up or completely stop between two sensors, creating outlier travel time samples. As such it is critical to screen those contaminated data first to minimize the bias of traffic state estimation.

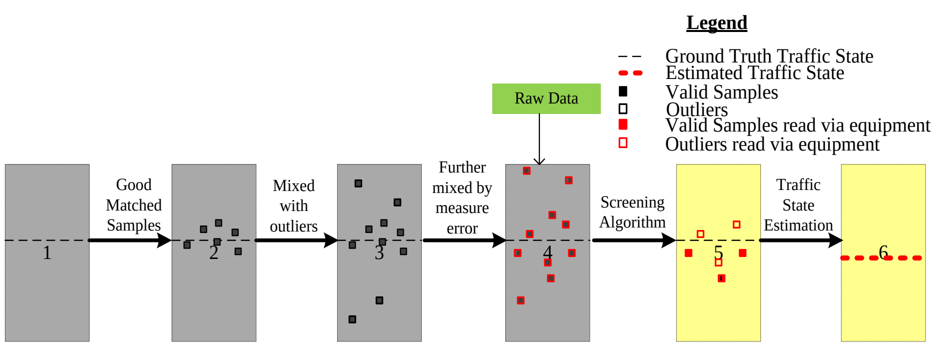


FIGURE 1 A generic process of traffic statement estimation via traffic sensors

Before describing the proposed screening algorithm, we first review some implemented screening algorithms in the past. Haghani et al. used multiple heuristic steps to process Bluetooth-based travel time samples. In their method, a set of 24-hour historical travel time samples are first processed to identify the travel speed distributions. It is then assumed that all valid travel speed samples during a time period would fall into the same average travel speed bins as did the historical data. Additional statistical techniques are then applied to further reduce suspicious outliers (*7*). Quayle et al. adopt a moving standard deviation to screen Bluetooth travel time samples falling outside of so called "cutoff limits"(*2*). In another screening algorithm to screen outliers, Boxel et al. conduct normality testing on travel time samples. If the data are normally distributed, it imply no outliers. Otherwise, they are considered outliers. If screening outliers was necessary, the stable travel speed is first calculated according to traffic density and the Greenshields model. Each sample is classified as an outlier (or not), based on its relationship to the stable travel speed, calculated using least quartile of squares (*8*). The screened Bluetooth data set were commonly used to estimate travel times and the proposed approaches can be divided into two categories:

1. Directly estimate the travel time with samples using various statistical models, such as the "seemingly unrelated Equations" (SURE) method (*9*), student distribution estimation(*2; 7; 10*), maximum likelihood estimation (*11-13*), the least quartile of squares method (*8; 14*), multiple variants of neural networks (*15-17*)
2. Bayesian estimation based on Kalman Filter Framework (*18; 19*) or Dynamic Bayesian Networks (*20*)

Upper bound and lower bound for valid travel time samples: the first step of data screening is to determine the upper bound and lower bound for valid samples given a traffic condition. According to the fundamental diagram of traffic flows, travel speed is correlated with the traffic volumes and we decide to adopt one of most widely adopted approach, the BPR function (*21*), to estimate the average travel speed under a given traffic volume. Obviously, not all travelers follow this speed in practice. Some move fast without meeting any forceful stop (e.g., traffic light) while some move slow but are stopped by all traffic control measures. Following the standard in traffic operations (*22*), we assume the 97.5 percentile drivers drive slower than the speed limit plus 10 miles per hour (MPH) while the 2.5 percentile of drivers drive faster than the speed limit minus 10 MPH. As such, the upper and lower bounds of non-stop travel times can be estimated accordingly. In addition, if the travel times are collected along arterials, the longest travel time should be further increased due to the control delays at intermediate intersections. Eq. [1~3] describes the travel time lower bound and upper bound given the speed limit, link length (and traffic signal timings if on arterials).

Without traffic signal control (Freeways): given road segment length L, free-flow travel speed u_f, time-dependent traffic volume v_t, the average travel speed u_d can be calculated based on the BPR function as:

$$u_d = \frac{u_f}{\left(1+0.15\left(\frac{v_t}{s}\right)^4\right)} \quad (1)$$

where s is the saturation rate (vehicles per hour per lane).

Accordingly, the upper bound $T_{UB}$and lower bound T_{LB} of the travel time samples are:

$$T_{LB} = \frac{L}{u_d+10}; T_{UB} = \frac{L}{u_d-10} \tag{2}$$

<u>*With traffic signal control (arterials)*</u>: given the number of intersections N, the longest waiting time at each intersection $w_i = C_i - g_{i,j}^{min}(i = 1,2, \dots, N)$ representing a vehicle arrives right after a minimum green expires, where C_i is the average cycle length at intersection i and $g_{i,j}^{min}$ is the minimum green of signal phase *j* along the subject link segment. As a result, the upper bound of travel time along signal-controlled arterials T'_{UB} are calculated as:

$$T'_{UB} = \frac{L}{u_d-10} + \sum_{i=1,2,\dots,N} w_i \tag{3}$$

Please note that Eq. (1-3) are relatively loose to avoid over screening valid samples.

4. Travel Time Estimation

Once the outliers are screened, we can use the remaining valid travel time samples to estimate link travel times. Recall that even the valid travel time samples still contain bias due to sensor's inherent measuring errors, it is necessary to pinpoint how the measuring errors are generated. As mentioned before, the timing error is the difference between the reported arriving time at road-side Wi-Fi sensors and its actual arriving time. The positioning error is the distance from the in-vehicle Wi-Fi device's actual location to the road-side Wi-Fi sensor. With a typical smartphone with known Wi-Fi MAC address at various distances from the customized road-side Wi-Fi sensor, we examine whether the known Wi-Fi MAC address can be captured and how strong the Wi-Fi signal is. It is considered that the minimal discoverable Wi-Fi signal strength is -90 dbi or above and therefore the location where a Wi-Fi MAC address is captured with the signal strength of -90 dbi is considered the sensing boundary of the road-side Wi-Fi sensors. FIG 2 shows the results from our field experiments. Three types of Wi-Fi antennas with various gains are used: 0 dbi (terminator), 1dbi and 5 dbi. From FIG 2, we can clearly tell that the higher gain the Wi-Fi antenna has, the larger area the Wi-Fi sensor can cover (i.e., the larger positioning errors) and we summarize the maximal positioning error of Wi-Fi MAC capturing with various speed limits and Wi-Fi antenna configurations in TABLE 2.

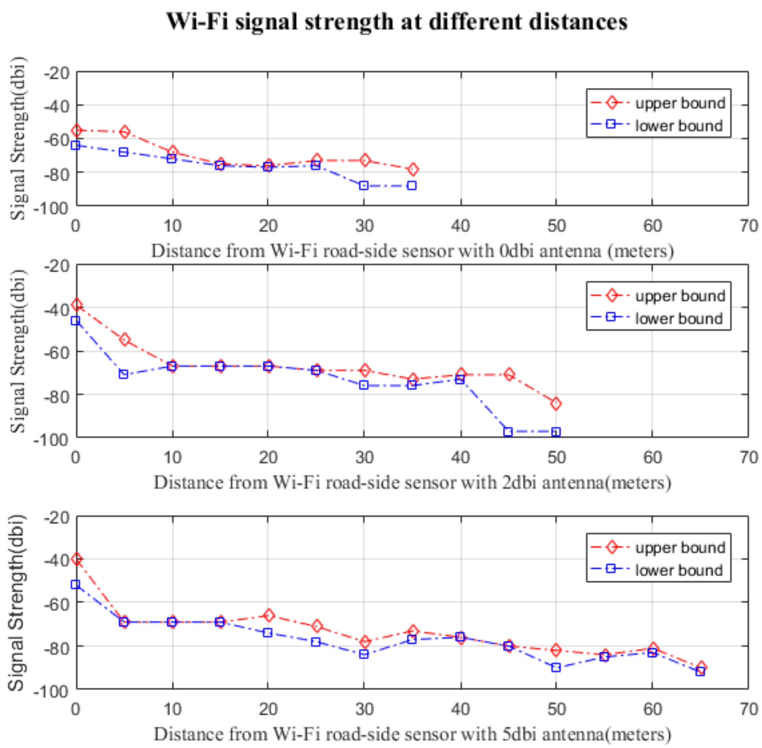


FIGURE 2 Signal strengths of Wi-Fi MAC address capturing at various distances

TABLE 2 Measuring error (in seconds) at one location in Wi-Fi MAC capturing technique

Wi-Fi Antenna/ Speed Limit	0 dbi	2 dbi	5 dbi
35 MPH	2.2s	3.2s	4.2s
45 MPH	1.7s	2.5s	3.2s
55 MPH	1.4s	2.0s	2.6s
65 MPH	1.2s	1.7s	2.2s
70 MPH	1.1s	1.6s	2.1s

Accordingly, the Wi-Fi sensor's timing error e_t can be calculated as:

$$e_t = \frac{R_i}{u_j} - l \quad (4)$$

where

R_i: the approximate radius of sensing area of the road-side Wi-Fi sensors under antenna configuration *i*;

l: the average latency from when a vehicle enters the sensing zone to when the Wi-Fi sensor discovers this vehicle which is reportedly up to one second (*4*);

u_j: The speed of vehicle *j*;

On the other hand, the positioning error e_d depends on vehicle *j*' speed and the radius of Wi-Fi sensing area. The maximum positioning error for one discovered vehicle can be calculated as:

$$e_d = R_i - u_j \times l \tag{5}$$

From Eq. (4-5), the total measuring error is the aggregation of two errors e_t, e_d and is formulated as:

$$e = e_t + e_d = \frac{R_i}{u_j} - l + R_i - u_j \times l = R_i \times \left(\frac{1}{u_j} + 1\right) - \left(1 + u_j\right)l = f(u_j) \tag{6}$$

Substitute u_j with $\frac{L}{t_j}$, we can get Eq. (7)

$$e = R_i \times \left(\frac{t_j}{L} + 1\right) - \left(1 + \frac{L}{t_j}\right)l = g(t_j) \tag{7}$$

where t_j is the segment travel time of vehicle *j*.

Since individual vehicle's speed (u_j) and link travel time (t_j) is random, Eq. (6) and (7) indicates that the total measuring error is also random and depends on the distribution of vehicle speed *u* or segment travel time *t*.

The distribution of *u* is commonly considered a Gaussian distribution (*22*) and so *t* may be modeled as an *Inverse Gaussian distribution* ($t \sim IG(\mu, \lambda)$) with probability density function:

$$f(t, \mu, \lambda) = \left[\frac{\lambda}{2\pi t^3}\right] * e^{-\frac{\lambda(t-\mu)^2}{2\mu^2 t}} \tag{8}$$

where: $\mu > 0$ (*mean*); $\lambda > 0$ (*shape parameter*) and they can be estimated as:

$$\hat{\mu} = \frac{\sum_{i=1}^{m} t_i}{m}; \hat{\lambda} = \frac{m}{\sum_{i=1}^{m}\left(\frac{1}{t_i} - \frac{1}{\hat{\mu}}\right)} \tag{9}$$

where t_i is one travel time sample (*i=1,2,…,m*).

In Eq. (9), the values of t_i are on an order of hundreds or even thousands of seconds. Therefore $\hat{l}$ will be a relatively large positive number if $m \geq 2$. Under this condition, the inverse Gaussian distribution can be approximated by the Gaussian distribution $\Phi(\mu', \lambda')$. In other words, travel time *t* can be modeled by a Gaussian distribution with mean travel time and variance estimated as:

$$\hat{\mu}' = \frac{\sum_{i=1}^{m} t_i}{m}, \hat{\sigma}'^2 = \frac{\sum_{i=1}^{m} \left(t_i - \hat{\mu}\right)^2}{m} \quad (10)$$

In case that sampling period is short or traffic volume is low, likely the number of valid travel time samples within a period k is small or zero. As a result, travel time estimation during k could be overwhelmed by individual sample's randomness. To mitigate this issue, we smooth the samples within two consecutive sampling periods, *k-1* and *k* to enhance the estimation accuracy and stability, formulated in (11):

$$\hat{\mu}_k' = \left(\frac{m_k}{\max\left(0.01, m_k + \max\left(0, \left(1 - \frac{T}{3600}\right)\right) m_{k-1}\right)} \right) \hat{\mu}_k + \left(1 - \frac{m_k}{\max\left(0.01, m_k + \max\left(0, \left(1 - \frac{T}{3600}\right)\right) m_{k-1}\right)} \right) \hat{\mu}_{k-1} \quad (11)$$

Where T is the sampling period (e.g., 900 seconds).

Eq. (11) imposes that the correlation between period *k-1* and *k* decreases when ΔT gets longer and two sampling periods become completely independent if ΔT is an hour or longer. In the meantime, the number of valid samples is linearly discounted according to the length of ΔT. If the sample size is zero during period *k*, then the estimated travel time during period *k-1* is used as its proxy.

Based on valid travel time samples and Eq. (7-11), the (Gaussian) distribution of Wi-Fi sensor measuring error can be inferred.

To take full advantage of the knowledge of Wi-Fi sensor's measuring error, we adopt a Kalman-Filter-based estimation framework initially proposed by Li and Souleyrette (*23*). The Kalman filter framework is widely used to estimate the true state of a linear dynamic system by minimizing the *Mean Squared Error* (MSE) (*24*). It is essentially a Bayesian estimating process except that the state space of the latent variables and observed variables are continuous and consistent with Gaussian distributions. Two steps are taken when a Kalman filter is applied: *Predict* and *Update*.

The predicting step is to estimate a system's true state according to its architecture and current state while the updating step is to correct the estimated results in the predict step using the observed samples (Bayesian). Mathematically, the Kalman filter approach assumes that the true travel time t_k during the period k is evolved from the true state t_{k-1} of period $k-1$ according to

$$t_k = F_k t_{k-1} + B_k u_k + w_k \tag{12}$$

During the period k, an observation (travel time sample) z_k of travel time is generated with:

$$z_k = H_k t_k + e_k \tag{13}$$

where:

F_k : Kalman state transition matrix during period k;

B_k : The control-input model of control vector u_k during period k;

H_k : The observation matrix in the Kalman filtering during period k;

e_k : The measuring error ($e_k \sim \Phi(0, R_k)$) during period k;

w_k : The system white noise ($w_k \sim \mathrm{F}(0, Q_k)$) during period k;

It is clear that all matrix and variables are of one dimension in this context.

Even though the traffic systems on freeways and on arterials are nonlinear in nature, various successful linearization efforts have been reported in the past, such as the first-order traffic flow model due to Newell. Under the Kalman filter framework, if travel time is collected on freeways, $B_k = 0$ since there are typically no control inputs. The ground truth travel times t_k, t_{k-1} from period *k-1* to period k is considered primarily affected by the volume changes and so we can derive theoretical travel time difference as $\Delta t = t_k - t_{k-1}$ based on Eq. (1). For simplicity, we further assume $t_k = \widehat{t_{k-1}} + \Delta t$ in the predicting step of Kalman filter, where $\widehat{t_{k-1}}$ is the estimated travel time based on Wi-Fi MAC address matching in period *k-1*.

In contrast, when travel time is collected on arterials, the ground truth travel time is affected not only by the volume changes but also by the signal timings at intersections between two Wi-Fi sensors. In essence, traffic signal control systems will reduce the road capacities along arterials and increase travel times. To address this feature, Eq. (1) is modified by replacing the saturation rate s with effective arterial road capacity c shown in Eq. (14). With Eq. (1) and Eq. (14), it is possible to predict the travel time $t_k = \widehat{t_{k-1}} + \Delta t$ in which both F_k and B_k are implicitly considered. Note that if the signal timing is not changed from period *k-1* to period *k*, then Δt is calculated the same as on freeways.

$$c = min\left(\frac{s \times g_i}{C_i}\right)(i = 1,2,\dots,N) \tag{14}$$

According to the *Traffic Engineering Handbook* (*22*), the 97.5th percentile and 2.5th percentile freeway travel speeds are typically about 14.7 feet/second (10 miles per hour) above and below the average speed, respectively. The corresponding link travel times can be calculated as $t_{2.5\%} = \frac{L}{\bar{u}+10\times1.47}$and $t_{97.5\%} = \frac{L}{\bar{u}-10\times1.47}$ (units are feet and seconds). Therefore, Q_k in w_k (the system noise) in Eq. (13) is calculated as:

$$s_t = \frac{(t_{97.5\%} - t_{2.5\%})}{2 * z_{0.025}} \tag{15}$$

where:

- $z_{0.025}$ can be obtained from the normal distribution table and its value is 1.96;

In Eq. (13), R_k (or e_k) is determined by the Wi-Fi sensor's total measurement errors discussed before and $H_k = 1$ as the travel time is directly measured. After all the parameters are determined, estimating travel time can be described below in Algorithm 1:

Lastly, smoothing method as in Eq. (11) is applied again to mitigate the issue of too small sample sizes within one sampling period.

Algorithm 1: Travel time estimation based on Kalman filter framework

Assume there are m_k valid travel time samples, z_k^i (*i=1, 2…,* m_k) collected during period *k*

Step 0: Initialization: when $k = 0$ (e.g., midnight), $t_0 = \frac{L}{u_f}$, w_k 's variance $Q_{k=0}$ is calculated with Eq. (15);

Step 1: for period *k* (k=1, 2 …), predict a priori estimation

Step 1.1: Predict the travel time with Eq. (1) and (14) as the priori estimation;

Step 1.2: estimate variance: $P_k^- = P_{k-1} + Q_k$;

Step 2: Calculate the Kalman gain: $K_k = P_k^- \left(P_k^- - R_k\right)^{-1}$(R_k is the variance of measure errors)

Step 3: Update

Step 3.1: update a posteriori state estimate: $\hat{t}_k = \hat{t}_k^- + K_k\left(z_k^i - \hat{t}_k^-\right)$

Step 3.2: covariance update: $P_k = \left(1 - K_k\right)P_k^-$

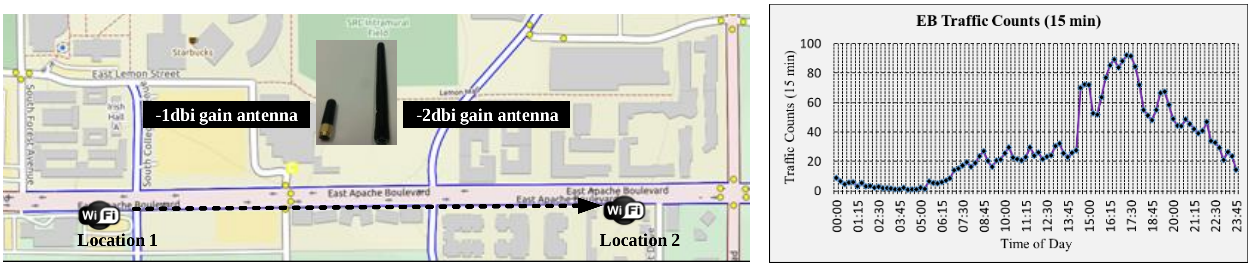
Step 4: if *i*< m_k , then *i=i+1* go to step 3; otherwise i=0, k=k+1 and go to Step 2

Step 5: if *i*< m_k , algorithm stops and the final travel time is retrieved as the estimate travel time during period *k*

5. Performance evaluation of Wi-Fi sensors in travel time estimations

5.1 Case One: travel time estimation in moderate traffic condition in Tempe, Arizona

A road segment along the Apache Blvd in Tempe, Arizona was selected to evaluate the performance of travel time estimation with captured Wi-Fi MAC addresses on October 2nd (Sunday), 2016. FIG.3-a shows two locations of road-side Wi-Fi MAC address collectors. The distance between two Wi-Fi road-side sensors is about 2,330 feet and the speed limit is 51 feet/second (35 miles per hour) and therefore the free-flow travel time between two collectors is around 45 seconds. There are several traffic signals between two Wi-Fi sensors and vehicles' actual travel times was longer than the free-flow travel time because of traffic. FIG. 3-b shows the traffic counts (east bound) for every 15 minutes. The background traffic along the Apache Blvd on that day was moderate and we observed that the green time was most assigned to the mainline. Therefore we ignored the impact of traffic signal control on the travel times for this case study.



a **b**

FIGURE 3 Locations of two Wi-Fi MAC collectors and two types of Wi-Fi antennas

The main objectives of this experiment include: examining the capability of capturing in-vehicle Wi-Fi MAC address of passing vehicles under various antenna configurations (-2dbi and -1dbi omni-direction antennas); examine the performance of proposed algorithms to filter outlier samples and estimate travel time over time.

The data collecting started at 10 AM on Oct-2nd, 2016 and lasted until 5:50 PM on the same day. The Wi-Fi antennas of both Wi-Fi MAC collector was first installed with the type of -3dbi omni-direction and then replaced with the -2dbi omni-direction antennas. According to FIG. 2, the maximum sensing radius for first antenna is 50 meters (164 feet) while the maximum sensing radius for the 2nd antenna is 30 meters (98 feet). The distributions of measuring errors under two types of antennas can be estimated according to Eq. (4) - (7) accordingly which will be used in estimating travel times with the Kalman filter framework later.

Wi-Fi MAC Capturing at one location (Vehicle presence data): The raw data collected by the Wi-Fi sensors were locally archived in a format of “Wi-Fi MAC address, epoch time” and then post processed. The first step of data processing was to identify how effective the Wi-Fi sensors could capture the nearby Wi-Fi MAC addresses. FIG. 4 shows the number of captured Wi-Fi MAC address in 5 minutes. Note that the antennas were replaced from -2dbi to -1dbi at 2 PM to reduce the sensing radius by 50%. From FIG. 4, it is clear that the number of captured Wi-Fi MAC addresses was only slightly reduced or not reduced at all after such changes. This makes sense because the scanning speed of Wi-Fi sensors is up to 5 Hertz (5 times per second) and it could almost capture all the Wi-Fi MAC addresses of passing objects within very small time windows. In contrast, typical Bluetooth-based travel time estimation must use large-gain antennas to capture sufficient Bluetooth MAC addresses. This finding suggests that it is promising to reduce the positioning and timing errors in Wi-Fi-based travel time estimation through special antenna configuration (e.g., low gain or directional) while the number of valid travel time samples is still enough for travel time estimation.

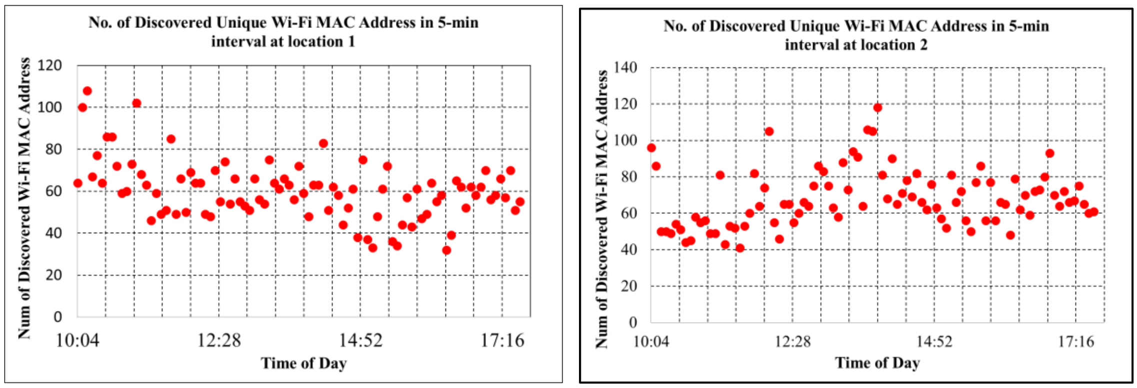


FIGURE 4 captured Wi-Fi MAC address at two locations (5-min interval)

__Travel Time Matching and travel time estimation__: the captured Wi-Fi MAC addresses at two locations are matched and screened according to the estimated upper and lower bounds. Then the travel time is estimated according to the proposed algorithm in this paper. FIG. 5 also shows some of the matched EB and WB travel time samples. Based on the several runs of floating vehicles during the experiment, the GPS travel time samples all fell into those 95% confidence intervals at the corresponding times.

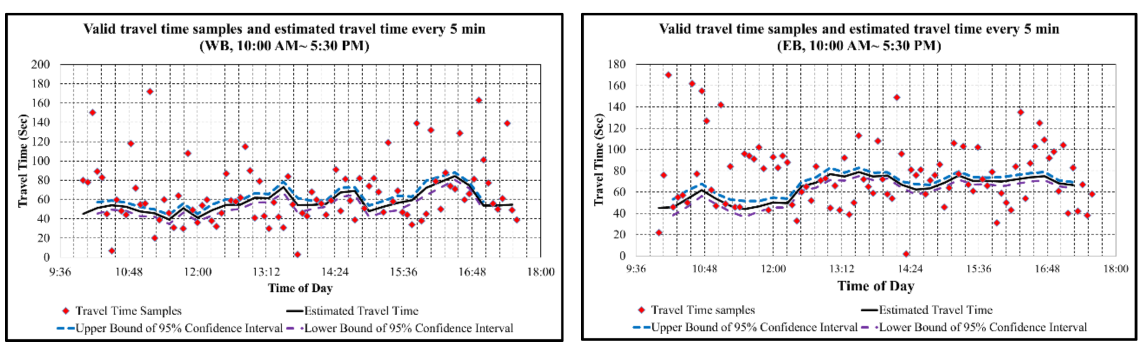


FIGURE 5 the number of valid travel time samples and estimated travel time over time

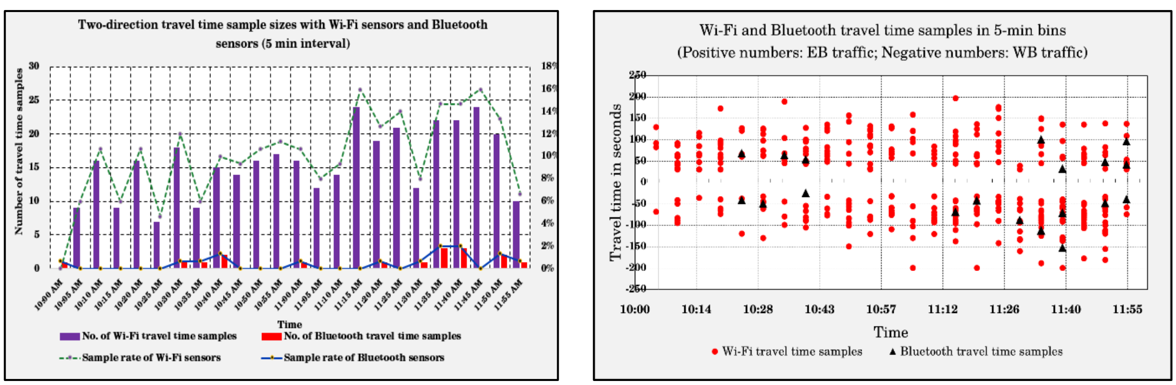
There are no data of turning movements at intermediate intersections between two Wi-Fi sensors. Nonetheless, it was observed that many vehicles turned into the campus of Arizona State University at intermediate intersections between Wi-Fi sensors and the travel time sample rate ranged from 10% to 30% approximately compared with the through traffic volumes.

5.2 Case 2: performance comparison between Wi-Fi sensors and Bluetooth sensors in Starkville, Mississippi

The purpose of this experiment it to conduct a cross comparison of the performance of the Wi-Fi sensors and Bluetooth sensors and it was conducted along the Highway 12 segment between Russel St. and Blackjack Rd. with the 45 MPH speed limit and 40 s free-flow travel time in Starkville, MS from 10:00 AM to 12: 00 PM (2 hours) on July 21st, 2017. Starkville has a relative low traffic during the summer and we count the two-direction arriving vehicles as 450 vehicles every 15 minutes during the experiment and the average observed speed is considerably slower than the speed limit. For the Wi-Fi sensors, we adopt 1dbi antennas and so the measuring error of travel time samples is about 2.0 s per location. For the Bluetooth sensors, we adopt a high-gain (5dbi) antennas in order to collect sufficient travel time samples and so its measuring error is about 10 s per location.

Sample Rate Comparison

The sample rate is critical in travel time estimation, especially in low-traffic areas. FIG. 6 reveals the numbers of matched travel time samples under Wi-Fi sensors and Bluetooth sensors and we can see that the Wi-Fi sensors considerably outperform the Bluetooth sensors in capturing and matching MAC addresses in passing vehicles under the same traffic conditions. The sample rate of Wi-Fi sensors ranges from 7% to 16% while that of Bluetooth sensors is in the range from 0% to 2%.

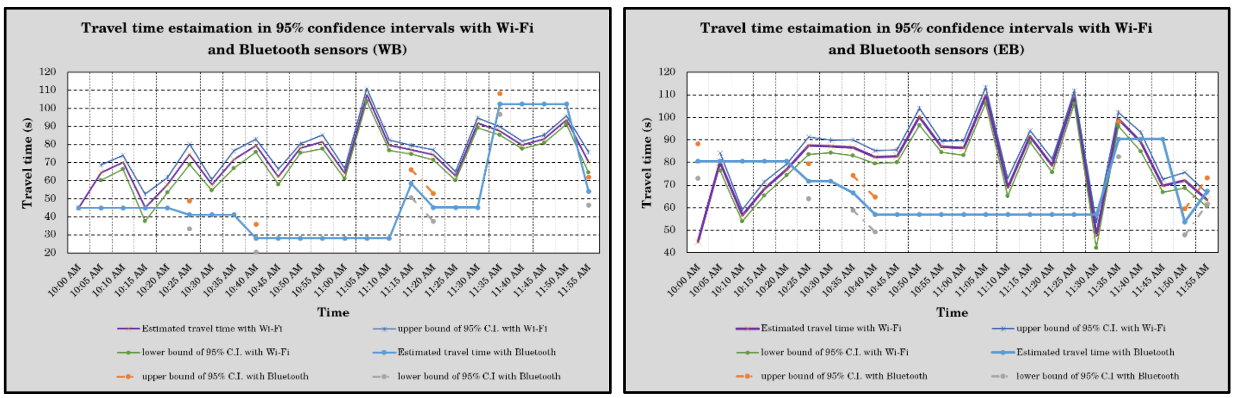


a b

FIGURE 6 Travel time sample rates of Wi-Fi sensors and Bluetooth sensors

Travel Time Estimation

With the proposed travel time estimation approach in this paper, the travel times as well as the boundary of 95% confidence interval are calculated every 5 minutes with the Bluetooth travel time samples and Wi-Fi travel time samples respectively. As revealed in FIG. 7, the Wi-Fi sensors are capable of capturing sufficient travel time samples in the low-traffic conditions and conduct good estimations. In contrast, the number of Bluetooth travel time samples are low in the low-traffic condition. Since there are no matched Bluetooth travel time samples in some time intervals, the proposed algorithm as in Eq. (11) will have to use the estimated travel time in the last sampling cycle (5 minutes in this context) and we can clearly tell the bias in the Bluetooth travel time estimation is considerably larger than the Wi-Fi travel time estimation.



a b

FIGURE 7 Travel time estimations with Bluetooth and Wi-Fi sensors

6 Conclusion

In this paper, we present a novel travel time estimation approach based on in-vehicle Wi-Fi MAC Capturing. To address two challenges in data processing: outlier screening and travel time estimation, we analyze the possible measuring errors of the proposed travel time estimation solution and then propose a Kalman-filter-based travel time estimation algorithm. To evaluate the performance of Wi-Fi sensors and to compare with the traditional Bluetooth sensors, two Wi-Fi MAC collectors and two Bluetooth sensors are built and deployed along two arterials in Arizona and Mississippi and to collect the Wi-Fi/Bluetooth MAC addresses in the passing vehicles. According to the experiments, we would like to draw the following conclusions:

1. the Wi-Fi sensors can adopt short-range antennas without creating the issue of low sample rates;

2. Depending on the vehicle speeds and Wi-Fi antenna configurations, the measuring errors of Wi-Fi sensors range from 1 s to 5 s per location;
3. The Wi-Fi sensors significantly outperform the Bluetooth sensors in capturing and matching travel time samples, especially in low-traffic areas.

In the future, we plan to continue exploring the potential of Wi-Fi sensors for other purposes, such as estimating control delays or queue lengths using the matched travel time samples if vehicles' presence times at locations of different Wi-Fi MAC collectors are synchronized with the high-resolution traffic signal events.

7 Acknowledgement

This work is partially funded by USDOT University Transportation Centers Contract DTRT13-G-UTC55/Sub-award UNR-14-60. The authors also thank Mr. James Sullivan and Ms. Clark Amanda of Mississippi Department of Transportation for their help and support in the field test in Starkville, Mississippi. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the official views or policies of the above organizations, nor do the contents constitute a standard, specification, or regulation of these organizations.

References

[1] Schrank, D., T. Lomax, and B. Eisele. 2011 Urban Mobility Report.In, Texas Transportation Institute, College Station, Texas, 2011. p. 59.
[2] Quayle, S. M., P. Koonce, D. Depencier, and D. M. Bullock. Arterial performance measures with media access control readers: Portland, Oregon, pilot study. *Transportation Research Record*, No. 2192, 2010, pp. 185-193.
[3] IEEE Standards Association. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications In, IEEE-SA Licensing and Contracts New York, NY 10016-5997, USA, 2012. p. 2793.
[4] Abedi, N., A. Bhaskar, and E. Chung. Bluetooth and Wi-Fi MAC address based crowd data collection and monitoring: benefits, challenges and enhancement. 2013.
[5] Porter, J. D., D. S. Kim, M. E. Magaña, P. Poocharoen, and C. A. G. Arriaga. Antenna characterization for Bluetooth-based travel time data collection. *Journal of Intelligent Transportation Systems,* Vol. 17, No. 2, 2013, pp. 142-151.
[6] Puckett, D. D., and M. J. Vickich. Bluetooth®-based travel time/speed measuring systems development.In, 2010.
[7] Haghani, A., M. Hamedi, K. F. Sadabadi, S. Young, and P. Tarnoff. Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record*, No. 2160, 2010, pp. 60-68.
[8] Van Boxel, D., W. H. Schneider, and C. Bakula. Innovative real-time methodology for detecting travel time outliers on interstate highways and urban arterials. *Transportation Research Record*, No. 2256, 2011, pp. 60-67.
[9] Martchouk, M., F. Mannering, and D. Bullock. Analysis of Freeway Travel Time Variability Using Bluetooth Detection. *Journal of Transportation Engineering,* Vol. 137, No. 10, 2011, pp. 697-704.
[10] Richardson, J. K., B. L. Smith, M. D. Fontaine, and S. M. Turner. Network stratification method by travel time variation. *Transportation Research Record*, No. 2256, 2011, pp. 1-9.
[11] Kwong, K., R. Kavaler, R. Rajagopal, and P. Varaiya. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 17, No. 6, 2009, pp. 586-606.
[12] Sanchez, R. O., C. Flores, R. Horowitz, R. Rajagopal, and P. Varaiya. Arterial travel time estimation based on vehicle re-identification using magnetic sensors: Performance analysis.In *14th IEEE International Intelligent Transportation Systems Conference, ITSC 2011, October 5, 2011 - October 7, 2011*, Institute of Electrical and Electronics Engineers Inc., Washington, DC, United states, 2011. pp. 997-1002.
[13] Sanchez, R. O., C. Flores, R. Horowitz, R. Rajagopal, and P. Varaiya. Vehicle re-identification using wireless magnetic sensors: Algorithm revision, modifications and performance analysis.In *Vehicular Electronics and Safety (ICVES), 2011 IEEE International Conference on*, 2011. pp. 226-231.
[14] Bhaskar, A., E. Chung, and A.-G. Dumont. Fusing Loop Detector and Probe Vehicle Data to Estimate Travel Time Statistics on Signalized Urban Networks. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 26, No. 6, 2011, pp. 433-450.
[15] Van Lint, J. W. C., and S. P. Hoogendoorn. A Robust and Efficient Method for Fusing Heterogeneous Data from Traffic Sensors on Freeways. *Computer-Aided Civil and Infrastructure Engineering,* Vol. 25, No. 8, 2010, pp. 596-612.

[16] Zeng, X., and Y. Zhang. Development of Recurrent Neural Network Considering Temporal-Spatial Input Dynamics for Freeway Travel Time Modeling. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 28, No. 5, 2013, pp. 359-371.
[17] Zhang, Y., and H. Ge. Freeway Travel Time Prediction Using Takagi–Sugeno–Kang Fuzzy Neural Network. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 28, No. 8, 2013, pp. 594-603.
[18] Barcelo, J., L. Montero, L. Marques, and C. Carmona. Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring. *Transportation Research Record*, No. 2175, 2010, pp. 19-27.
[19] Qiu, Z., P. Cheng, J. Jin, and B. Ran. Cellular probe technology applied in advanced traveller information system. *World Review of Intermodal Transportation Research*, Vol. 2, No. 2, 2009, pp. 247-260.
[20] Hofleitner, A., R. Herring, and A. Bayen. Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. *Transportation Research Part B: Methodological*, Vol. 46, No. 9, 2012, pp. 1097-1122.
[21] Bureau of Public Roads. Traffic Assignment Manual.In, Dept. of Commerce, Urban Planning Division, Washington D.C., 1964.
[22] Roess, R., E. Prassas, and W. McShane. *Traffic Engineering (4th Edition)*. Prentice Hall, New York City, 2010.
[23] Li, P., and R. R. Souleyrette. A Generic Approach to Estimate Freeway Traffic Time Using Vehicle ID- Matching Technologies. *Computer–Aided Civil and Infrastructure Engineering*, Vol. 31, No. 5, 2016, pp. 351-365.
[24] Kalman, R. E. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, Vol. 82, No. 1, 1960, pp. 35-45.