

**Final Technical Report  
TNW2008-05**

**Research Project Agreement No. 61-4230**

# **Improved Positioning for Fleet Management and Traveler Information**

Dan J. Dailey

Department of Electrical Engineering  
University of Washington  
Seattle, WA 98195

A report prepared for

**Transportation Northwest (TransNow)**  
University of Washington  
135 More Hall, Box 352700  
Seattle, Washington 98195-2700

in cooperation with

**U.S. Department of Transportation**  
Federal Highway Administration

March 2008

## TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO. <b>TNW2008-05</b>	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.	
4. TITLE AND SUBTITLE <b>Improved Positioning For Fleet Management And Traveler Information</b>		5. REPORT DATE <b>March 2008</b>	
		6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) <b>Dan Dailey</b>		8. PERFORMING ORGANIZATION REPORT NO. <b>TNW2008-05</b>	
9. PERFORMING ORGANIZATION NAME AND ADDRESS <b>Transportation Northwest Regional Center X (TransNow) Box 352700, 129 More Hall University of Washington Seattle, WA 98195-2700</b>		10. WORK UNIT NO.	
		11. CONTRACT OR GRANT NO. <b>DTRS99-G-0010</b>	
12. SPONSORING AGENCY NAME AND ADDRESS <b>United States Department of Transportation Office of the Secretary of Transportation 400 Seventh St. S.W. Washington, D.C. 20590</b>		13. TYPE OF REPORT AND PERIOD COVERED <b>Final Research Report</b>	
		14. SPONSORING AGENCY CODE	
15. SUPPLEMENTARY NOTES <b>This study was conducted in cooperation with the University of Washington and the US Department of Transportation.</b>			
ABSTRACT <b>This proposal investigates the possibilities for transit management to take advantage of the ever increasing wireless/Wi-Fi infrastructure in urban areas. In previous projects with Metro and Sound Transit, we have evaluated the use of GPS positioning for fleet management and traveler information. While the GPS positioning worked well for both busses on the regional freeways and trains on regional tracks, it deteriorated when the vehicles entered the dense urban core. This is caused by tall buildings or other structures occluding the GPS satellites. This deviation can be sufficient to place the vehicle on the wrong street and disrupt fleet management and traveler information. An additional source of positioning information is becoming more prevalent by the day, the Wi-Fi hotspot. These act as beacons at known locations. This beacon-like position can be integrated with GPS though optimal filtering.</b>			
17. KEY WORDS <b>Traffic, Pedestrian, GPS</b>		18. DISTRIBUTION STATEMENT	
19. SECURITY CLASSIF. (of this report) <b>None</b>	20. SECURITY CLASSIF. (of this page) <b>None</b>	21. NO. OF PAGES <b>39</b>	22. PRICE

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the views or policies of the Washington State Transportation Commission, Department of Transportation, or Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

# Contents

<b>Executive Summary</b>	<b>iv</b>
<b>1 Introduction/Problem Statement</b>	<b>1</b>
<b>2 Research Objectives</b>	<b>3</b>
<b>3 Background</b>	<b>5</b>
<b>4 Analysis</b>	<b>7</b>
4.1 Wi-fi Framework . . . . .	7
4.2 Seattle Wireless access points . . . . .	12
4.3 Data Collection and Results . . . . .	12
4.4 Transit Vehicle Data . . . . .	22
4.5 Data Fusion Theory . . . . .	24
4.5.1 Data Fusion . . . . .	27
<b>5 Conclusion and Recommendation</b>	<b>28</b>
5.1 Conclusion . . . . .	28
5.2 Recommendation . . . . .	29
<b>A Determination of filter parameters</b>	<b>32</b>

# List of Figures

1.1	Roadway as a line, in state plane coordinates, and GPS positions as points in downtown Seattle. . . . .	2
3.1	Architecture for multi-modal use of Wi-fi to augment positioning. . . . .	6
4.1	Normalized power as vehicle passes Wi-fi Station at constant velocity. . . . .	8
4.2	The effect of motion and one second sampling period. . . . .	10
4.3	The effect of motion and one second sampling period. . . . .	11
4.4	University District Wi-fi coverage. . . . .	15
4.5	Data taken from the King County kc_streets.E00 file. . . . .	16
4.6	University avenue digital map data and example Wi-fi access point locations from the Wigle Database. . . . .	16
4.7	University Avenue pictographic map from Seattle GIS department. . . . .	17
4.8	Example location data for a measurement set. (Note that the tick marks on the horizontal axis are 100 feet apart.) . . . . .	18
4.9	Model fit to data from February 4, 2008 with external high gain antenna on the left and January 29,2008 on the right with built-in antenna for each access point. . . . .	20
4.10	Model fit to data from February 4, 2008 with external high gain antenna on the left and January 29,2008 with built-in antenna for each access point. . . . .	21
4.11	University avenue digital map data with corrected Wi-fi access point locations and Transit Data from 2/4/2008. . . . .	23

# Executive Summary

This work investigates the possibilities for transit management to take advantage of the ever increasing wireless/Wi-Fi infrastructure in urban areas.

The first area of interest involves the Wi-Fi on selected Sound Transit routes. At present the Metro AVL system is based upon dead reckoning using an odometry reading with digital maps and a set of signposts as reference points. At present when the vehicles take alternative routes, such as snow days, the tracking breaks down due to the vehicles not passing assigned signposts. However, with the installation of many Wi-Fi access points throughout the metropolitan area the combination of Wi-Fi access on the vehicle and a mapping of Wi-Fi access points can effectively provide a very high density of virtual signposts that will allow tracking even when the vehicles are off the standard route due to weather conditions, emergencies or driver error.

The second area is the soon to be GPS equipped Sound Transit busses operated by Pierce transit and Community transit. Soon to be installed passenger counting equipment has GPS capabilities included in the package. Under a previous project with Sound Transit we have demonstrated tracking prediction capabilities along the freeways; however, in the dense heart of the urban area the tracking and prediction capabilities degraded substantially due to the inability to reliably obtain satellite ranges. Using existing Wi-Fi sites as beacons would vastly improve the tracking in the urban area.

# Section 1

## Introduction/Problem Statement

This work investigates the possibilities for transit management to take advantage of the ever increasing wireless/Wi-Fi infrastructure in urban areas. In previous projects with Metro and Sound Transit we have evaluated the use of GPS positioning for fleet management and traveler information. While the GPS positioning worked well for busses on the regional freeways and trains on regional tracks it deteriorated when the vehicles entered the dense urban core. This is caused by occlusion of the GPS satellites by tall buildings or other structures. Figure 1.1 shows the results from earlier tests where the vehicle track deviates from the actual street locations.[1] This deviation can be sufficient to place the vehicle on the wrong street and disrupt fleet management and traveler information. An additional source of positioning information is becoming more prevalent by the day, the Wi-Fi access point. These act as beacons at known locations, when integrated with odometry or GPS through optimal filtering.

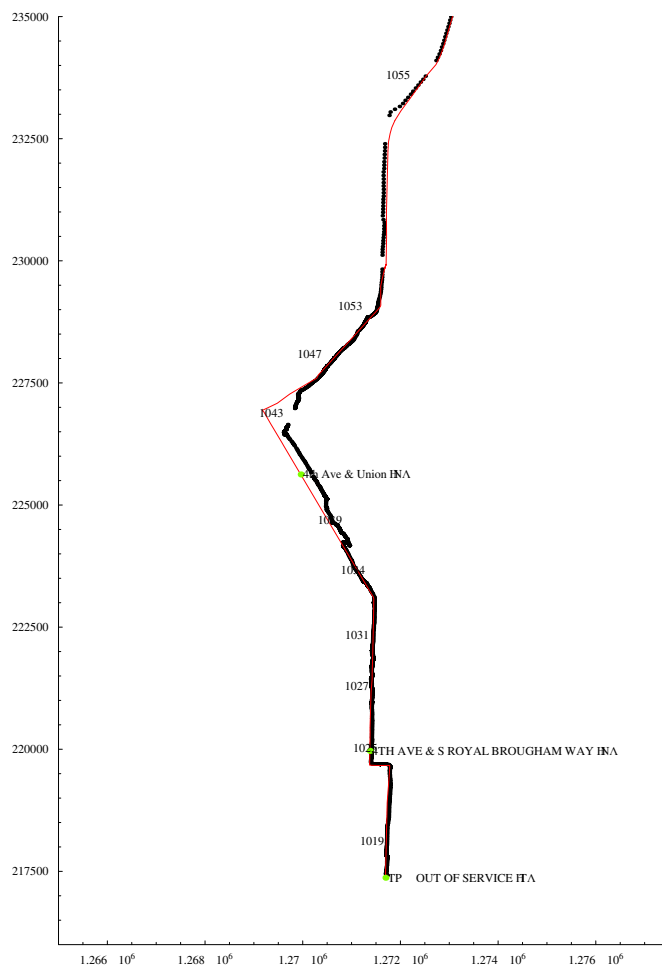


Figure 1.1: Roadway as a line, in state plane coordinates, and GPS positions as points in downtown Seattle.



## Section 2

# Research Objectives

This work has at least three areas of impact on the Puget Sound regional transit activity, one affecting vehicles operated by Metro Transit, one on vehicles operated by Pierce and Community Transit, and one that impacts traffic management and traveler information.

The first area of impact involves the effect of Wi-Fi on selected Sound Transit routes operated by Metro Transit. At present the Metro AVL system is based upon dead reckoning using odometry with digital maps and a set of signposts as reference points to reset the odometers. Again at present when the vehicles take alternative routes, such as on snow days or driver error, the tracking breaks down because the vehicles do not pass assigned signposts. However, with the installation of many Wi-Fi access points throughout the metropolitan area these access points can act as auxiliary sign posts. The combination of Wi-Fi access on the vehicle with a mapping of Wi-Fi access points, can provide a very high density of virtual signposts that will allow tracking, even when the vehicles are off the standard route due to weather conditions, emergencies or driver error.

The second area of impact is the soon to be GPS equipped Sound Transit coaches operated by Pierce Transit and Community Transit. Soon to be installed passenger counting equipment has GPS capabilities included in the package. Under a previous project with Sound Transit we have demonstrated tracking and prediction capabilities along the freeways; however, in the dense heart of the urban area the tracking and prediction capabilities degraded substantially due to the inability to reliably obtain satellite ranges. Using

existing Wi-Fi sites as beacons will improve the tracking in the urban areas and improve real-time traveler information.

A third impact is on the associated use of the transit data as probe vehicle data. Improvements in positioning make the use of busses as virtual probe vehicles more accurate. There are a number of areas in Seattle where: (1) the busses travel, (2) there is no infrastructure for traffic parameter measurement and (3) there is a desire to have traffic management information, such as SR99. This contribution extends beyond the usefulness to transit and benefits traffic management.

The academic content comes on two fronts. The first is developing algorithms and implementations to integrate GPS position fixes with beacon information and digital maps in an optimal filtering framework. This will extend the work found in [1] that created and implemented a multi-agency, multi-county tracking and prediction system.

The second academic contribution is in geographic reasoning. The list of access points is a slowly varying one that needs to be kept current and correct. In order to use the access points reliably for geo-location several issues must be addressed: (1) validating each access point as it is encountered (2) identifying new access points with the proper location, (3) placing the access points on the correct trip in the transit spatial schedule, (4) differentiating off-route behavior from access point changes, and (5) identifying and updating information on access points that have changed. This slowly changing database problem will be solved by recording metrics of performance, such as range to access point verses distance into trip, and using multiple hypothesis testing in real-time to update the database.

## Section 3

# Background

Several projects have been performed with Sound Transit to evaluate positioning, prediction, and communications technologies, and the results published in [1, 2]. These projects were done in the context of past work on optimal prediction of transit departure [3], providing departure information wirelessly [4], transit vehicle as probes [5, 6], and using this probe information for traffic management [7]. Figure 3.1 shows the addition of Wi-Fi receivers as positioning beacons to the framework used successfully in [1] and [2].

Wi-Fi access points transmit a unique identifier as well as some naming information when polled by a Wi-Fi enabled device. Past work has used time of flight measures for positioning in 2D, but this requires customized hardware and modifications to the Media Access Control (MAC) layer of the IEEE 802.11 protocols. Other work has used signal strength measurements to estimate distance based on the  $1/r^2$  radial reduction in power. In the work most closely related to that proposed here, a mapping of signal strength as a function of position, usually inside a building, is first created a priori and then future location estimates are made using knowledge of this relationship. One advantage to this project is the knowledge that the vehicles will be on the roadway, thus limiting the positioning to finding the linear location along a roadway when the location of the roadway is well known.

This report describes several steps in the research: (1) obtain either a mapping or map a set of Wi-Fi access points that lie on a large number of transit routes, (2) place laptops on vehicles that travel chosen routes to obtain signal strength measures for various access points, (3) implement al-

gorithms for positioning in Matlab and use initial vehicle data to test and evaluate algorithms, (4) develop theory for combining Wi-Fi location data with GPS reports in an optimal filter framework, and (5) evaluate the affect of additional position information on tracking and prediction capabilities.

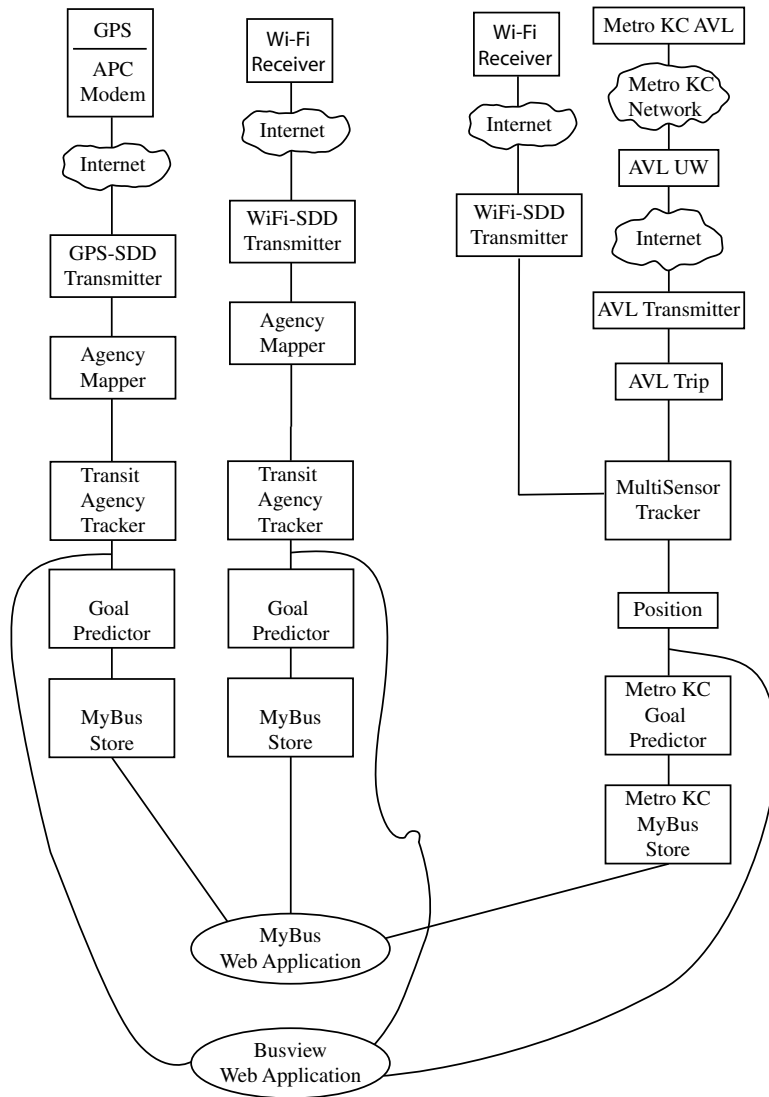


Figure 3.1: Architecture for multi-modal use of Wi-fi to augment positioning.

# Section 4

## Analysis

### 4.1 Wi-fi Framework

The model for signal strength as a function of range from a Wi-fi station is based on free space propagation model (sometimes called the Friis transmission equation [8, 9, 10])

$$\frac{P_R}{P_T} = G_R G_T \left( \frac{\lambda}{4\pi d} \right)^2$$

where  $P_R$  is the power available at the receiving antenna and  $P_T$  is the power supplied to the source antenna;  $G_R$  is the receiver antenna gain and  $G_T$  is the transmitter antenna gain;  $\lambda$  is the carrier wavelength;  $d$  is the transmitter-receiver distance.

The vehicle is traveling on the roadway

$$x(t) = x_0 + v(t)t + \frac{1}{2}a(t)t^2$$

where  $x$  is along the roadway.

The Wi-fi station is located off of the roadway in the  $y$  direction. The minimum distance between the station and the roadway is  $d_0$  located at  $x'$ . This distance between receiver and transmitter is

$$d(t) = \sqrt{(x(t) - x')^2 + d_0^2}$$

$$R_{ij}(t) = C_{ij} \left( \frac{\lambda}{4\pi \sqrt{(x(t) - x')^2 + d_0^2}} \right)^2$$

where  $R_{ij}$  is the power ratio for the  $i^{th}$  station for the  $j^{th}$  receiver and  $C_{ij}$  are the combined constants for the  $i^{th}$  station for the  $j^{th}$  receiver.

Channel 6, in the middle of the Wi-fi band, is at a frequency of 2437MHz, and  $\lambda = c/f$ , where  $c$  is the speed of light, and  $f$  is the frequency. Speed of light is 299,792,458 m/s therefore the wavelength for Channel 6 is 12.3 cm.

Figure 4.1 shows the shape and relative magnitude of the normalized power ratio for differing distances (two, five, and 10 meters) perpendicular to the roadway as a vehicle passes a access point at a constant velocity. This is calibrated by making measurements along the roadway.

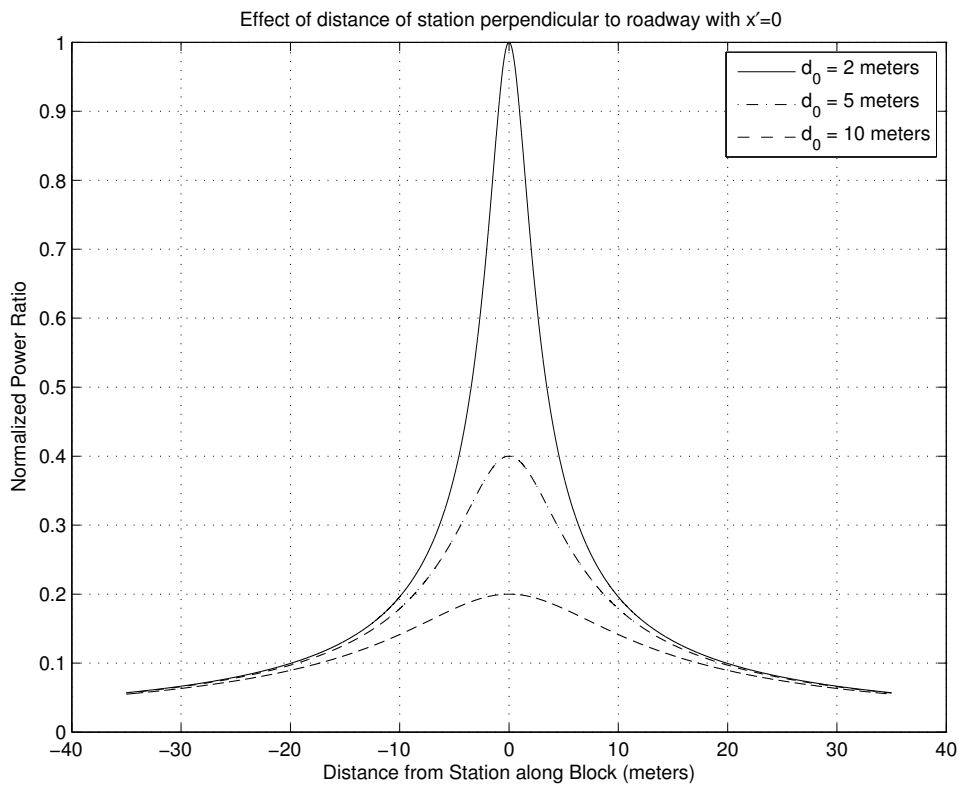


Figure 4.1: Normalized power as vehicle passes Wi-fi Station at constant velocity.

Many applications use the National Marine Electronics Association (NMEA) standard for communication with the a GPS receiver that has an update rate of once per second. The networking hardware and software used in this project does active scanning. It sends a probe request about once a second and records the responses.

In the analysis presented here the sample rate is once per second and the change in the response function based on the vehicle speed can be seen in the changes in Figure 4.2. The vehicle is modeled as moving down the street or block where the access point is 200 feet from the closest approach of the vehicle. The horizontal axis is the range to the access point as the vehicle moves along the block. The figures present the statistical effect on the expected received signal power of sampling once per second with a random start time for vehicles ranging in speed from 5 miles per hour to 60 miles per hour. The conclusion from these calculations is that the positioning activity is best done at relatively slow speed like those found on city streets.

In Figure 4.3 the motion effect on the signal strength as a function of time is presented. Again this calculation suggests that slow vehicle speeds will provide the best positioning information.

The conclusions from this analysis are: (1) using access points near the street will provide the best localizing information and (2) this approach works best from relatively slow moving vehicles like those on streets and arterials and is unlikely to work on freeways. The next step is to develop a methodology to define "passing" an access point so as to generate a location measurement.

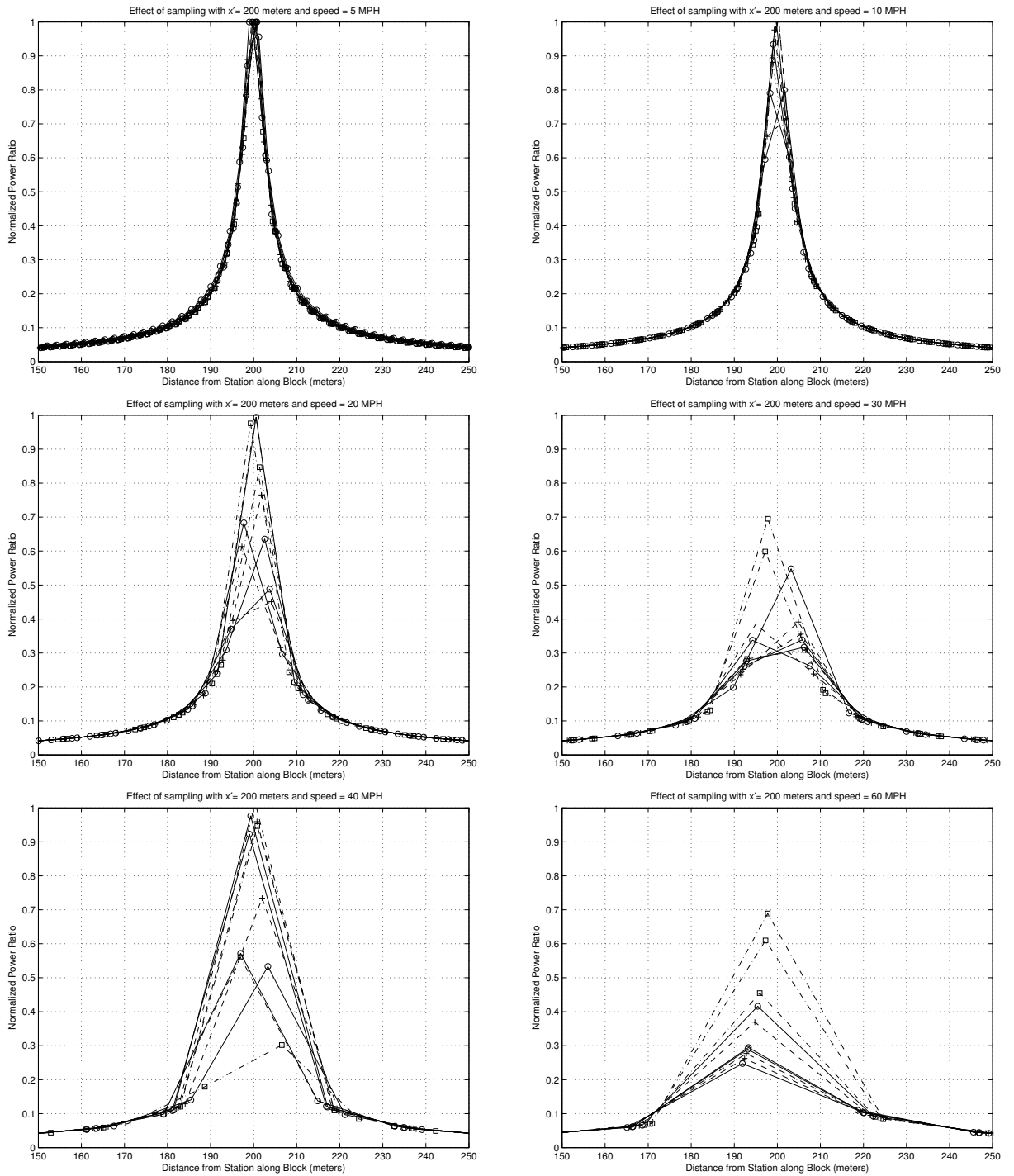


Figure 4.2: The effect of motion and one second sampling period.



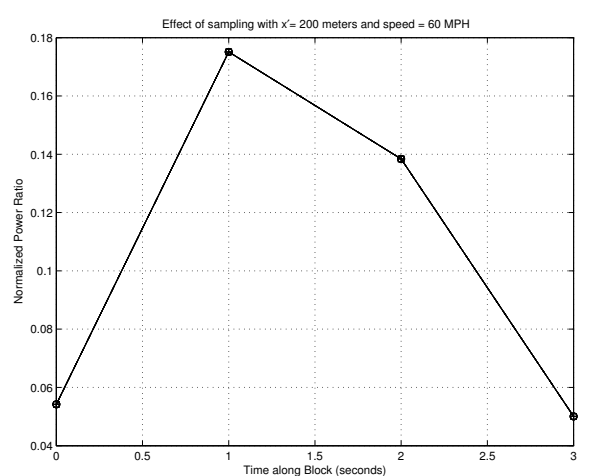
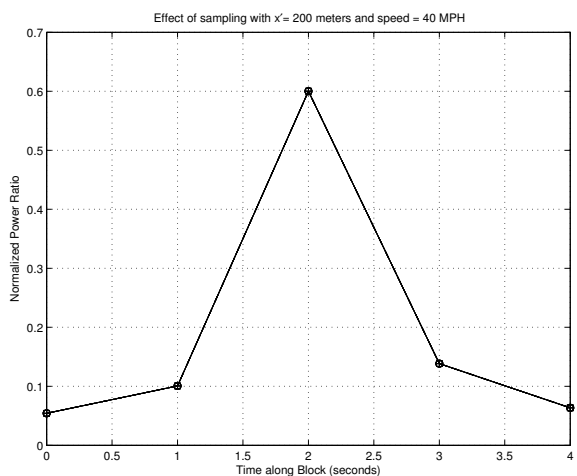
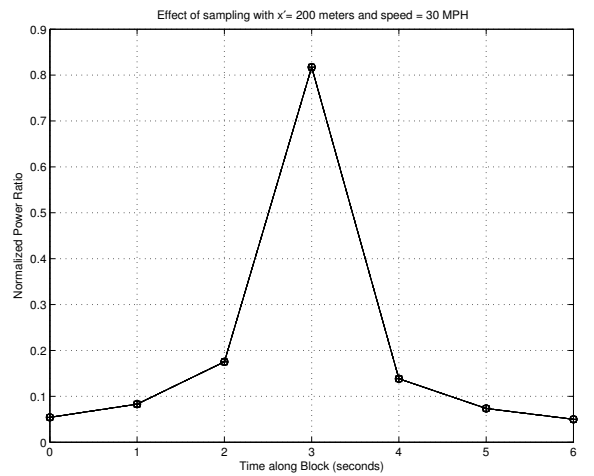
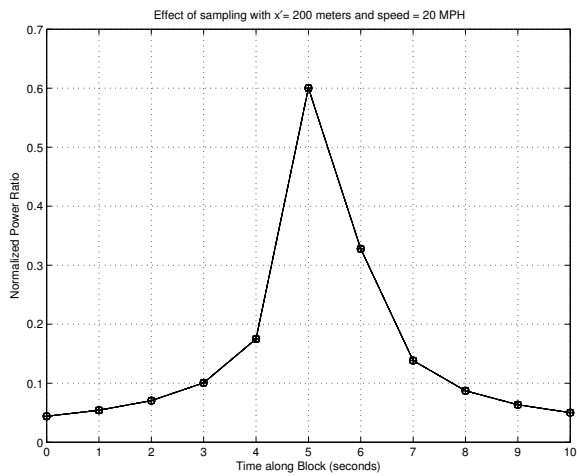
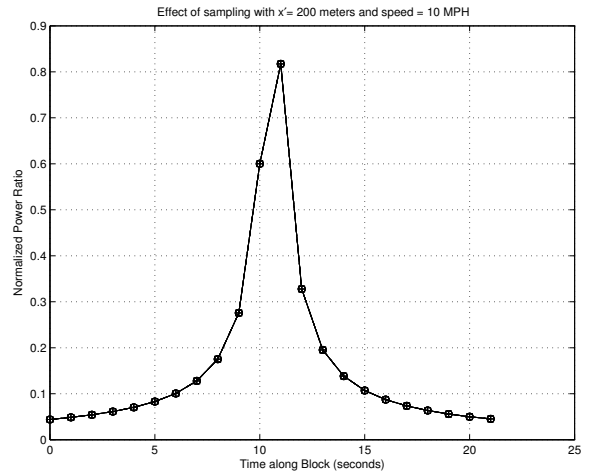
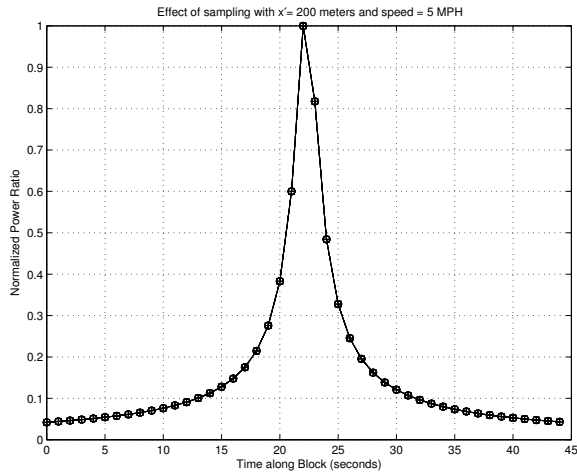


Figure 4.3: The effect of motion and one second sampling period.

## 4.2 Seattle Wireless access points

The City of Seattle has deployed Wi-fi access points along a number of popular streets, in particular in the University district. The reasoning behind this deployment is:

“The City of Seattle is providing free wireless Internet access in the Columbia City and the University District business districts. The City’s Wi-Fi pilot project also includes four downtown Seattle parks: Occidental, Freeway, Westlake and Victor Steinbrueck, as well as the City Hall lobby area. This is a pilot project. Users can log-in using “seattlewifi” for the SSID. In 2006, Seattlewifi served 14,400 different users.”

The area of deployment in the University district is shown in Figure 4.4. “In the University District, coverage currently runs along University Way (The Ave) from 40th Street NE to approximately midway between 45th St. NE and 47th”

“The goals of the City’s Wi-Fi pilot project are: to attract more customers to local business districts, support small businesses, encourage the use of public parks and facilities, and enable more citizens to access City services online.”

The set of Wi-fi access points deployed by the city of Seattle have a “service set identifier” (SSID) or network name of “seattlewifi” that can be used along with the MAC address to identify a particular device at a particular location. This set of access points is the basis for this positioning work.

## 4.3 Data Collection and Results

To identify calibration constants and verify the ability to properly locate the time and location of passage events the following experiments are done. Each of the experiments records data on a Microsoft Windows XP based computer for post processing. The hardware and software consist of: an IBM Thinkpad with builtin Wi-Fi antenna, a Proxim pcmcia Wi-fi card, a 6db gain external antenna, a USB GPS receiver, and Netstumbler software

found at <http://www.netstumbler.com/>.

The experiments include:

1. Driving past fixed access point at various speeds
2. Driving a bus route with access points at various speeds
3. Following a bus performing revenue service over the same routes

An underlying digital base map/data set is needed both to validate measurements and locations, and to model the motion of the vehicles along the roadway. The base map from Metro-King County used for the MyBus and Busview projects was obtained in an Arcview .E00 interchange format. It was converted to an ESRI coverage format using the ESRI IMPORT71 utility. The coverage is represented by an ESRI shapefile which is a published binary format for geographical information. The Matlab program was used to load and crop the data to the region of the University District. Since the eventual goal is to track transit vehicles the cropping is based on using two time points from the Metro-King County Automatic Vehicle Location systems: timepoint 5506 for 15 AV NE & NE 65 ST located at (1276177,250004) in Washington State Plane Coordinates, and timepoint 5952 for BROOKLYN AV NE & NE CAMPUS PKWY (NB) located at (1275427,242883) in Washington State Plane Coordinates. The resulting data set is plotted in Figure 4.5, where the circles represent the nodes to be used to represent the vehicle planned path through the region.

The next need was an underlying database of installed Wi-fi access points and their WA state plane locations. Based on suggestions from [11] the Wigle database (<http://www.wigle.net>) was selected as the source for Wi-fi access point locations. The set of stations extracted from the wigle database, in html format, was identified by selecting a 0.1 mile radius around the intersection of 15th Avenue NE 42nd Street. The resulting html was parsed using the perl script `html2txt.pl` by Matti J. Karki ([mjk@iki.fi](mailto:mjk@iki.fi)) with active perl 5.8.820, and the tuple (SSID, MacID, latitude, longitude) extracted.

The data from Wigle is recorded in latitude and longitude but in order to perform distance measures easily the (lat, lon) is projected onto Washington

North State Plane coordinates using a Lambert conformal conic projection [12].

Example University District locations and media access control addresses (MAC addresses) as presented in the Wigle data base are shown in Figure 4.8. A pictorial representation of the location of the access points, as obtained from the GIS section of the city of Seattle, is shown in 4.7

In the course of making the measurements a vehicle equipped with the networking equipment is driven north and south on the University Avenue. The output of the Netstumbler software include the latitude and longitude as estimated by the GPS received at the time of each sampling of active access points.

The data from the measurements included position, signal strength, noise level, and signal-to-noise ratio. Wi-fi performance levels are often reported as a signal-to-noise ratio. The signal-to-noise ratio threshold is used in identifying useful signals for network connections and values below 20 are considered “low.”

A vehicle was driven north and south on University Avenue recording both GPS positions and probing for Wi-fi access points once per second. Access points with the SSID of “seattlewifi” were extracted from this data. The MAC address for the “seattlewifi” access point was used with the wigle database to extract the location, as recorded in wigle, for the access point. This location is shown on the left of Figures 4.9 through 4.10. In the right of Figures 4.9 through 4.10 the signal-to-noise ratio for each of the access points is plotted against the range between the GPS reported position and the wigle location of the “seattlewifi” access point.

An example set of measurement points for a northbound and south bound trip are shown in Figure 4.8.

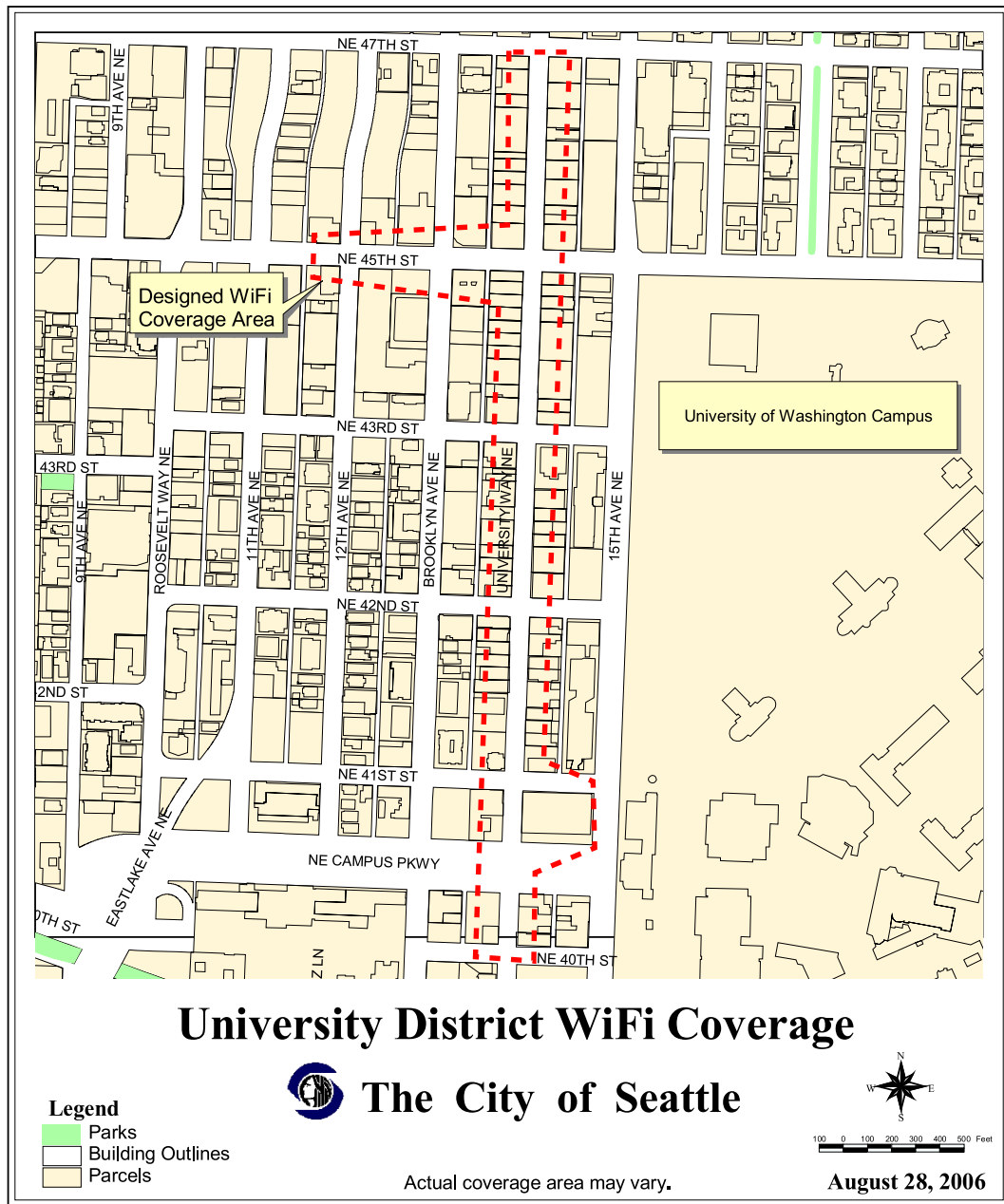


Figure 4.4: University District Wi-fi coverage.

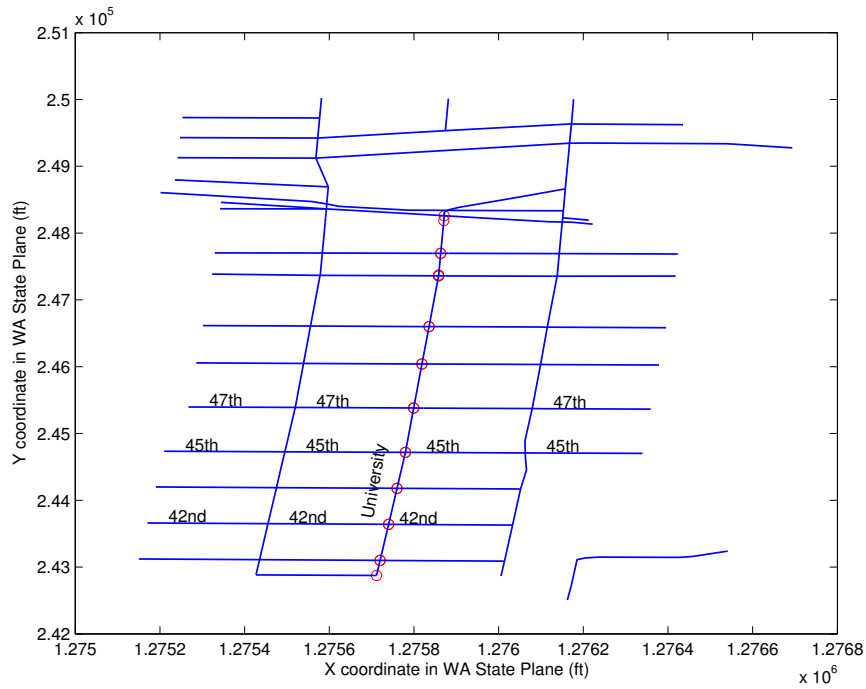


Figure 4.5: Data taken from the King County kc\_streets.E00 file.

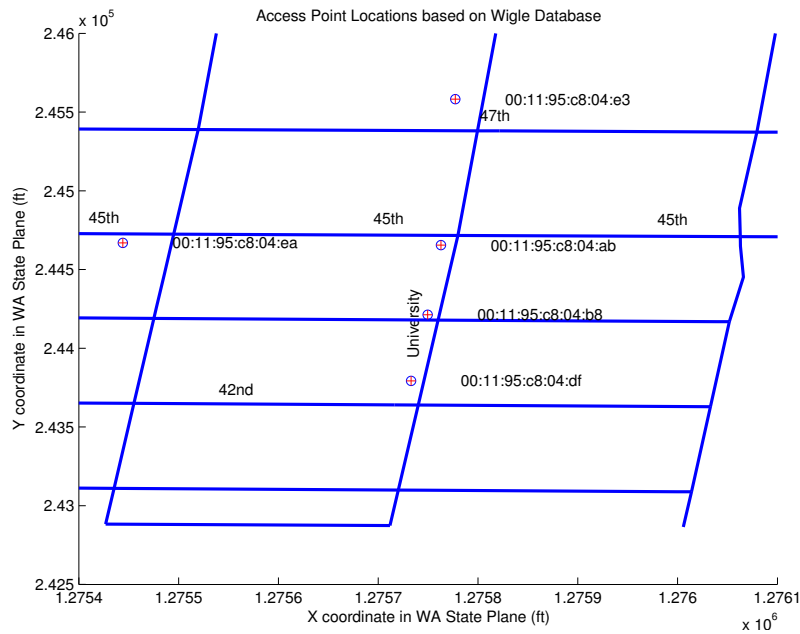


Figure 4.6: University avenue digital map data and example Wi-fi access point locations from the Wigle Database.



Figure 4.7: University Avenue pictographic map from Seattle GIS department.

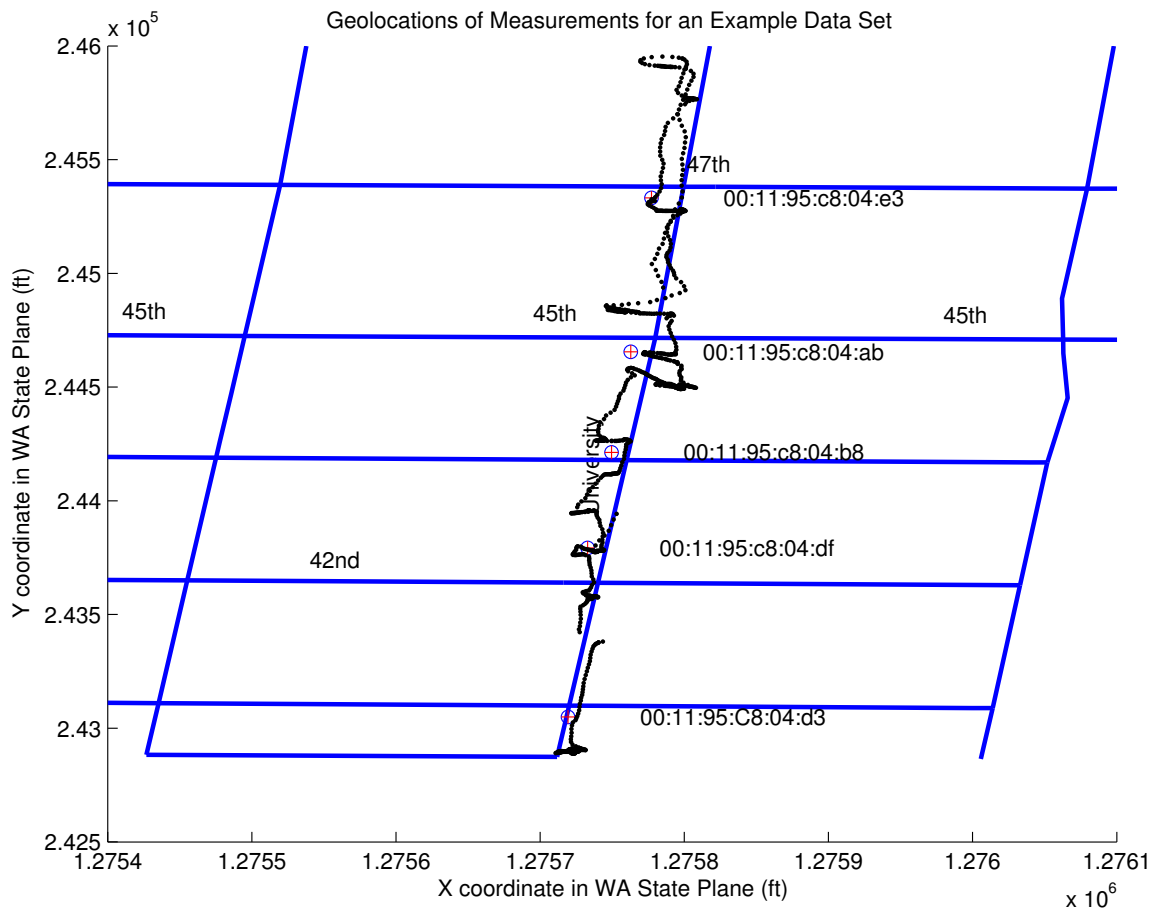


Figure 4.8: Example location data for a measurement set. (Note that the tick marks on the horizontal axis are 100 feet apart.)



The simple range model developed in section 4.1 was augmented with two additional parameters in order to fit the data. One to estimate the offset north/south of the GPS fix to the wigle locations estimate ( $c$ ) and a noise floor parameter ( $b$ ),

$$SNR = \frac{A}{\sqrt{(R - c)^2 + d^2}} + b \quad (4.1)$$

where  $A$  is the normalized amplitude and  $R$  is the range. This model is nonlinearly least squares fit to the data.

The result of the fitting process is shown in Figures 4.9 and 4.10 for the set of active stations along University Avenue. Data was obtained both with a high gain antenna and with the built-in antenna. The model fit is much better with the high gain antenna, shown in the left of the figures, both because the data density is much higher and the signal-to-noise ratio appears more quantized with the less sensitive internal antenna. In both cases a passing event can clearly be defined. In the case of the high gain antenna, for this set of access points, simply using the first observed signal-to-noise ratio value over 60 might work rather well as a surrogate to the full fitting procedure.

Since the University Avenue is oriented primarily North-South the parameter  $c$  effectively represents the north-south deviation of the measurement fit from the wigle database's geo-location of the access points. If the access point location data base is corrected using this parameter the new access point locations now agree with the geolocation suggested pictorially by the City of Seattle GIS department. This can be confirmed by comparing Figures 4.8 and fig:SeaGIS.

This section provides analysis and support for the notion that a passing event can be defined at the level of a few feet. The data collected using the high gain antenna are superior to those using the built-in antenna for defining a passing event but both work at some level. A surrogate measurement using a threshold on the signal-to-noise ratio,  $SNR > 60$  for these access points may even be adequate to define a passing event to tens of feet when using the high gain external antenna. The accuracy of the positioning using the existing King County AVL system based on odometry is approximately 500ft,

so a passing event on the order of feet is a vast improvement in the individual positioning estimate.

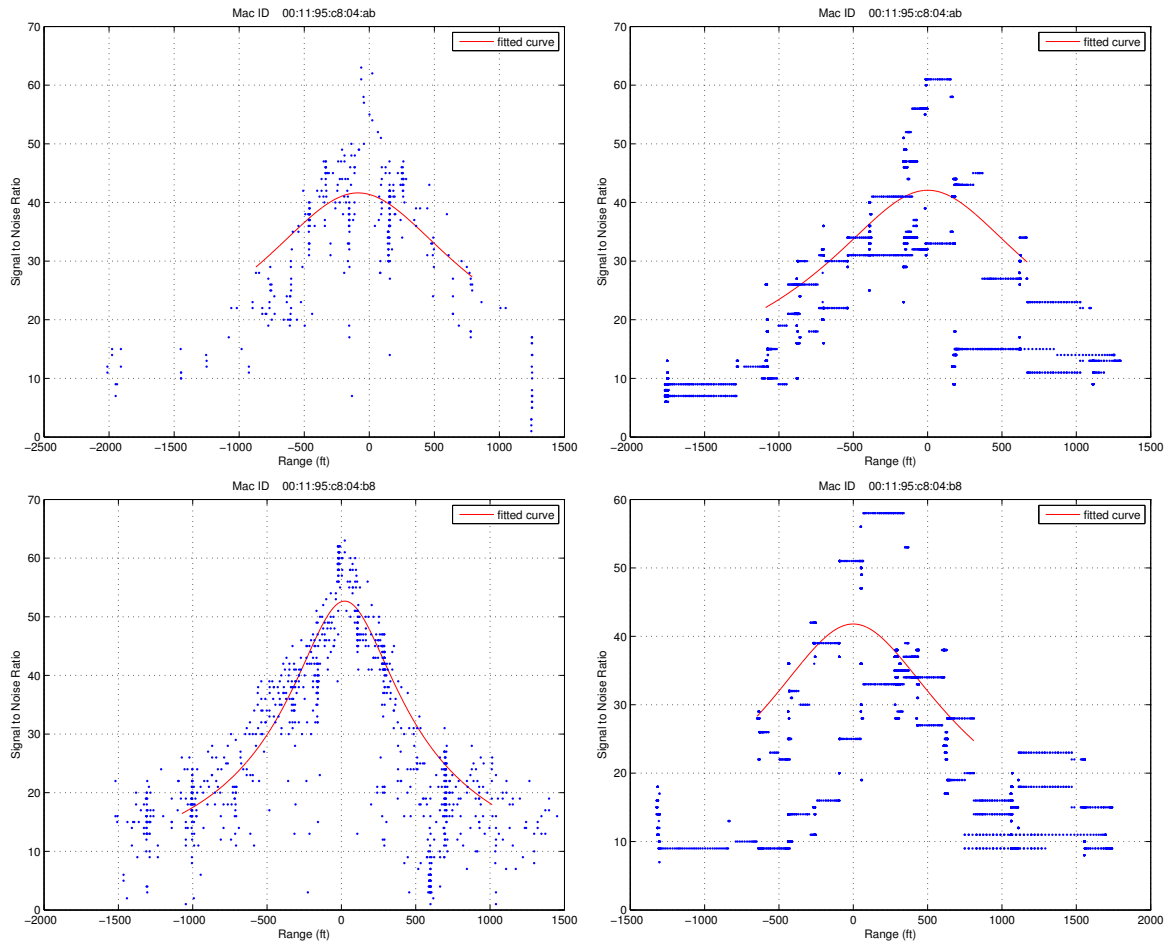


Figure 4.9: Model fit to data from February 4, 2008 with external high gain antenna on the left and January 29,2008 on the right with built-in antenna for each access point.

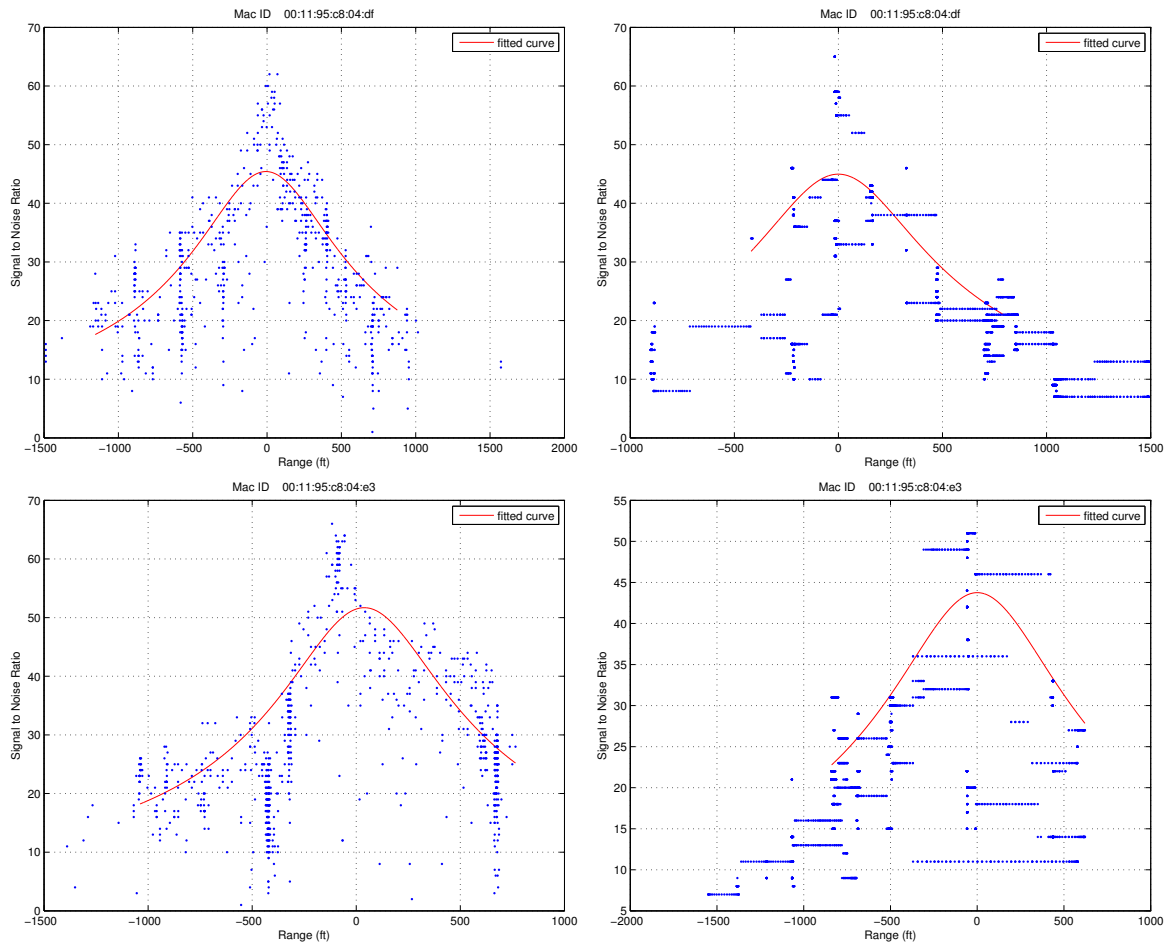


Figure 4.10: Model fit to data from February 4, 2008 with external high gain antenna on the left and January 29, 2008 with built-in antenna for each access point.

## 4.4 Transit Vehicle Data

During the collection of Wi-Fi and GPS data several transit vehicles were followed. Examples are shown in Table 4.1. To obtain the quantitative values for the locations of the vehicles as a function of time data was collected from the system shown in Figure 3.1. A data replay component was built to allow the raw transit data to be augmented using the Kalman filter positioning components built for the Mybus and Busview applications. In particular the AvlSelector, AvlTracker2, AvlTranslator, and Position components were instantiated as shown in figure 3.1 and the output of the Position component was recorded for use in demonstration combining the two forms of positioning.

Table 4.1: Example tracking trips.

Vehicle ID	Service Route	Description	Approximate Time	Date
2415		Northbound	5:33 pm	1/29/2008
2774	71E	Southbound	5:41 pm	1/29/2008
2811	74	Northbound	5:11 pm	2/4/2008
2790	72	Southbound	5:17 pm	2/4/2008

For the February 4 data two vehicles were followed while recording the Wi-Fi signal to noise ratio, the GPS position and the Odometry based positioning. Figure 4.11 shows the vehicle position data for two vehicles in the University district that day, as well as the location of the access points. It is clear from the proximity of the odometry measurements and the access point location that they will be complementary data sources.

The mechanisms and applications to extract detailed transit data for use in building positioning algorithms that perform data fusion on various types of data is the main contribution of this portion of the work. It is now possible to extract raw transit from Metro, archive it and at a later time post process it to obtain positions in both geodetic and state plane coordinates as well as the covariance matrix from the Kalman filter used to estimate the position.

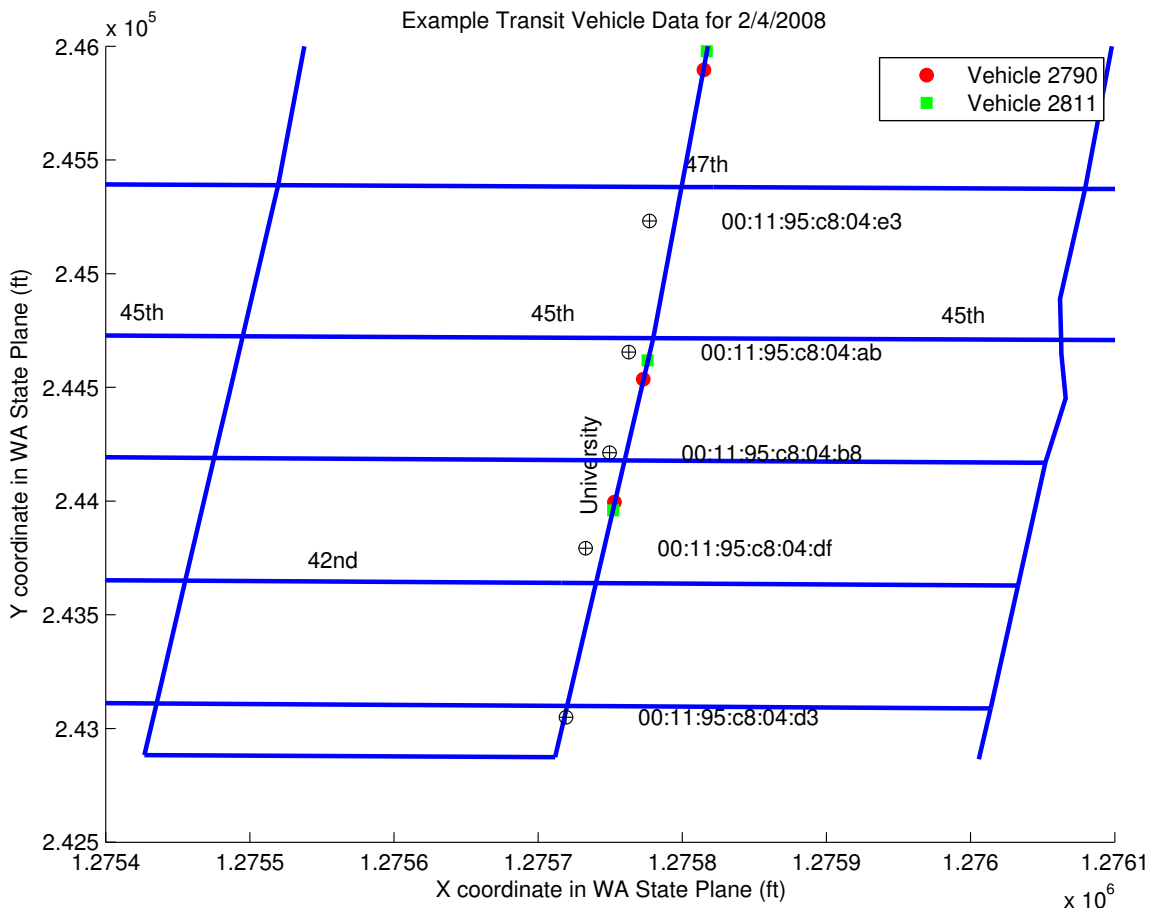


Figure 4.11: University avenue digital map data with corrected Wi-fi access point locations and Transit Data from 2/4/2008.

## 4.5 Data Fusion Theory

In past work we use Kalman filtering concepts to estimate vehicle positions based on odometry based AVL data. In this section the basic theory to perform data fusion for the two measurement types, odometry and Wi-fi beacons is developed. In order to implement a Kalman filter the following must be specified:

- a state-space,
- a measurement model,
- a state transition model,
- an initialization procedure.

Once these items are specified, one may employ any one of a number of implementations of the Kalman filter/smoothen equations, (see [13], [14], [15] or [16]) to transform a sequence of measurements into a sequence of vehicle state estimates.

To represent the instantaneous state of a vehicle we select a three dimensional state vector:  $x$  is distance into pattern, or along the roadway in meters,  $v$  is speed in meters per second, and  $a$  is acceleration in meters per second squared. The vehicle state vector is

$$X = \begin{bmatrix} x \\ v \\ a \end{bmatrix}.$$

We assume a second order dynamics for the state transition. This is defined by the first order system of linear stochastic differential equations

$$\begin{aligned} dx &= v dt \\ dv &= a dt \\ da &= dw. \end{aligned} \tag{4.2}$$

Here  $dt$  is the differential of time and  $dw$  is the differential of Brownian motion representing randomness in vehicle acceleration. By the definition of

Brownian motion (see Chapter 3, Section 5 of [14]), the variance is

$$E(dw^2) = q^2 dt$$

where  $q^2$  is a model parameter. In the absence of a measurement correction the variance of acceleration grows linearly with time. We selected a value for  $q^2$  of  $(264 \text{ ft}/\text{min}^2)^2/\text{min} = (3 \text{ mph}/\text{min})^2/\text{min}$  using the method described in appendix (A).

The differential equations (4.2), are written in vector form as

$$dX = F X dt + G dw$$

where

$$F = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad G = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Integrating over a time interval  $(t, t + \delta t)$ , (See the text following Theorem 7.1 of [14] for a discussion of stochastic integration) we obtain the state transition model

$$X(t + \delta t) = \Phi(\delta t) X(t) + W(\delta t)$$

where  $X(t)$  and  $X(t + \delta t)$  denote vehicle state values at times  $t$  and  $t + \delta t$  respectively,

$$\Phi(\delta t) = \exp(F \delta t) = \begin{pmatrix} 1 & \delta t & \delta t^2/2 \\ 0 & 1 & \delta t \\ 0 & 0 & 1 \end{pmatrix}$$

is the “transition matrix”, and where the accumulated error  $W(\delta t)$  has covariance

$$Q(\delta t) = \begin{pmatrix} \delta t^5/20 & \delta t^4/8 & \delta t^3/6 \\ \delta t^4/8 & \delta t^3/3 & \delta t^2/2 \\ \delta t^3/6 & \delta t^2/2 & \delta t \end{pmatrix} q^2.$$

In order to run a Kalman filter/smoothen we need an initialization procedure, a method for computing an initial value for the state vector and its associated error covariance matrix. These initial values,  $\hat{X}_0$  and  $P_0$ , are based on the initial measurement  $z_0$  (at time  $t_0$ ) and measurement variance  $R$ . We set  $\hat{X}_0 = (z_0 \ 0 \ 0)$  and set

$$P_0 = \begin{pmatrix} R & 0 & 0 \\ 0 & (30\text{mph})^2 & 0 \\ 0 & 0 & (16\text{mph}/\text{min})^2 \end{pmatrix}.$$

The initialization of the covariance above is specified by a number of “hard-coded” parameters.

There are two types of measurements used, odometry and Wi-fi beacons. These measurements have approximately independent errors and a Kalman filter is run for both measurement sets. The optimal state vector is constructed by combining them using a method described in [17].

The first measurement, the odometry measurement,  $z$ , is an estimate of the vehicle’s distance into trip, and this measurement model is given by

$$z_1 = H_o X + \epsilon_o = x + \epsilon_o.$$

Here  $H_o = (1 \ 0 \ 0)$  is the “measurement matrix” and  $\epsilon$  denotes a random measurement error, assumed to have a Normal distribution with variance  $R_o$ . The variance is treated as a model parameter with nominal value of  $R_o = (500\text{ft})^2$  where the value for  $R$  was determined as described in subsection (A). The second measurement  $z_2$ , the Wi-fi beacon location, is treated similarly but with a different parameter,

$$z_2 = H_w X + \epsilon_w = x + \epsilon_w. \tag{4.3}$$

and  $\epsilon_w$  denotes a random measurement error, assumed to have a Normal distribution, with variance  $R_w$  that is determined by experiment.

Using the notation from section 5.2.3 of [18], the Kalman filter equations are:

Predicted state

$$\hat{X}(k+1|k) = \Phi(k)\hat{X}(k|k)$$

Predicted measurement

$$\hat{Z}(k+1|k) = H\hat{X}(k+1|k)$$

State Prediction Covariance

$$P(k+1|k) = \Phi(k)P(k|k)\Phi(k)^T + Q(k)$$

Measurement Prediction covariance

$$S(k+1) = HP(k+1|k)H^T P(k+1|k) + R$$



Filter gain update

$$W(k+1) = P(k+1)H^T S(k+1)^{-1}$$

State update

$$\hat{X}(k+1|k+1) = \hat{X}(k+1|k) + W(k+1)(Z(k+1) - \hat{Z}(k+1|k))$$

Covariance update

$$P(k+1|k+1) = \Phi(k+1|k) - W(k+1)S(k+1)W(k+1)^T$$

This is the Kalman framework used to do positioning with odometry data, the next step is to perform data fusion with both odometry data and Wi-Fi base data.

### 4.5.1 Data Fusion

Optimal positioning is done using data fusion to make a position estimate in a filtering frame work. The two data to be fused are positions from odometry  $\hat{X}_1$  and Wi-fi passing events  $\hat{X}_2$ . The data to be used is based on observations from each data stream (1) odometry,  $z_1$ , and (2) Wi-fi beacons,  $z_2$ . An optimal fused state estimate is shown in Section 8.6 of [17] to be

$$\begin{aligned} \hat{X}(k|k) = & P(k|k)\{P(k|k-1)^{-1}\hat{X}(k|k-1) \\ & + [P_2(k|k)^{-1}\hat{X}_2(k|k) - P_2(k|k-1)^{-1}\hat{X}_2(k|k-1)] \\ & + [P_1(k|k)^{-1}\hat{X}_1(k|k) - P_1(k|k-1)^{-1}\hat{X}_1(k|k-1)]\} \end{aligned}$$

and the covariance update is,

$$\begin{aligned} P(k|k) = & \{P(k|k-1)^{-1} + P_2(k|k)^{-1} - P_2(k|k-1)^{-1} \\ & + P_1(k|k)^{-1} - P_1(k|k-1)^{-1}\}. \end{aligned}$$

It is this set of equations that are used to fuse the two data streams.

## Section 5

# Conclusion and Recommendation

### 5.1 Conclusion

This report develops a framework to evaluate and implement the use of Wi-Fi access points as beacons for positioning. It develops a simple range based signal strength model for the Wi-Fi stations. The conclusions from the numerical simulation are: (1) using access points near the street will provide the best localizing information and (2) this approach works best from relatively slow moving vehicles like those on streets and arterials and is unlikely to work on freeways.

The report develops a simple range based estimate for the signal strength from an access point. It presents the results from least squares fitting the model to several sets of data. Data was taken over several days and with two sets of hardware. The data collected using the high gain antenna were found superior to those using the built-in antenna for defining a passing event but both work at some level.

Initially the access point positions were obtained from the Wigle database, but it was concluded that this publicly available set of data was not accurate enough to perform vehicle positioning.

An additional conclusion was that a surrogate measurement using a threshold on the signal-to-noise ratio (SNR),  $SNR > 60$  for these access points may even be adequate to define a passing event to tens of feet when using the high gain external antenna. This can also provide a mechanism for surveying a

block of transit work for access point to later be used for positioning.

The accuracy of the positioning using the existing King County AVL system, based on odometry, is approximately 500ft, so a passing event on the order of feet is a vast improvement in the individual positioning estimate.

The work performed on this project created a software framework to take raw transit data and compute positioning data, including the covariance measurements of the position estimates, that can then be fused with the Wi-Fi passing data in a Kalman filter framework.

Finally, the report describes a theoretical framework for a Kalman filter to fuse Wi-Fi positioning with other positioning system outputs. Due to time constraints the mathematical framework has yet to be implemented in software.

## **5.2 Recommendation**

The data fusion algorithm needs to be implemented in software to demonstrate the improved positioning capabilities using Wi-Fi access point positioning.

Additional field tests are needed with improved hardware so that the data fusion algorithm can be quantitatively evaluated.

# Bibliography

- [1] D.J. Dailey, F.W. Cathey, and S.D. Maclean. Design and realization of a multi-modal/multi-agency transit management and information system. In *Proceedings of the Informs Annual Meeting, 2004*, October 2004.
- [2] F.W. Cathey, D.J. Dailey, and S.D. Maclean. Data integration and web applications for a multi-modal/multi-agency transit management and information system. In *Proceedings of the Seventh Annual World Congress on Intelligent Transport Systems*, October 2004. Nagoya, Japan.
- [3] F.W. Cathey and D.J. Dailey. A Prescription for Transit Arrival/Departure Prediction using AVL Data. *Transportation Research C*, 11C(3-4):241–264, 2003. June-Aug.
- [4] S.D. Maclean and D.J. Dailey. Wireless Internet Access to Real-Time Transit Information . *Transportation Research Record*, 1791:92–98, 2002.
- [5] F.W. Cathey and D.J. Dailey. Estimating corridor travel time by using transit vehicles as probes. *Transportation Research Record*, 1855:60–65, 2003.
- [6] F.W. Cathey and D.J. Dailey. Transit vehicles as traffic probe sensors,. *Transportation Research Record*, 1804:23–30, 2002.
- [7] F.W. Cathey and D.J. Dailey. Implementation of traffic management using transit probes. In *Proceedings of the IEEE 8th International Conference on Intelligent Transportation Systems*, September 2005. Vienna, Austria.
- [8] K. Feher. *Wireless Digital Communications*. Prentice Hall, 1995.

- [9] F. Lassabe, P. Canalda, P. Chatonnay, and F. Spies. A Friis-based Calibrated Model for WiFi Terminals Positioning. In *Proceedings of the Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM05)*, June 2005.
- [10] J.D. Parsons and J.G. Gardiner. *Mobile Communications Systems*. Halsted Press, John Wiley, 1989.
- [11] J. Hightower, A. LaMarca, and I. Smith. Practical Lessons from Place Lab . *IEEE Pervasive Computing*, 5(3):32–39, 2006.
- [12] J. E. Stern. *State Plane Coordinate System of 1983*. NOAA Manual NOS NGS 5, January 1989.
- [13] B. M. Bell. The Marginal Likelihood for Parameters in a Discrete Gauss-Markov Process. *IEEE Transactions on Signal Processing*, 48(3), March 2000.
- [14] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.
- [15] F. Tung H. E. Rauch and C. T. Striebel. Maximum Likelihood Estimates of Linear Dynamic Systems. *American Institute of Aeronautics and Astronautics Journal*, 3:1445–1450, August 1965.
- [16] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, Inc, 1979.
- [17] Y. Bar-Shalom and X.R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. YBS Publishing, 1995.
- [18] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. Wiley, 2000.

# Appendix A

## Determination of filter parameters

As mentioned above, our measurement and process models depend on two parameters: the measurement variance,  $R$ , and the rate of change of the variance of process noise,  $q^2$ . Using the method of “maximum marginal likelihood” we obtained “optimal” values for these parameters in a number of experiments with different measurement sequences. Our goal was not to perform an exhaustive statistical analysis, but rather just to find some representative values that would give reasonable filter performance. We observed that the values obtained in each experiment were roughly the same, ie, fluctuated around  $R = (500 \text{ ft})^2$  and  $q^2 = (3 \text{ mph/min})^2/\text{min}$ .

Although the method of maximum marginal likelihood for parameter estimation is well known to statisticians, the fact that it can be used effectively for estimating parameters in the setting of Kalman-Bucy filters is not widely reported. This method was first proposed in [13] where an effective procedure is provided for evaluating the marginal likelihood function. We briefly describe the theory.

Let  $\mathbf{Z}$  and  $\mathbf{X}$  denote vector-valued random variables representing a measurement sequence and a corresponding sequence of state vectors, and let  $\xi$  denote the parameter vector  $(R, q^2)$ . A formula is given in [13] (Eq.4) for the joint probability density function  $p_{\mathbf{Z}, \mathbf{X}}(z, x; \xi)$  in terms of the measurement and process models like those described above. (The symbols  $z$  and  $x$  in this context denote real-valued vectors and usage is not to be confused with that

in the preceding section.) The cited reference provides an algorithm (Algorithm 4) which, when given a measurement vector  $z$  and parameter vector  $\xi$ , simultaneously computes the maximum likelihood estimate for the state vector sequence (the Kalman smoother estimate)

$$x_{z,\xi}^* = \operatorname{argmax}_x \{p_{\mathbf{X}|\mathbf{Z}}(x|z; \xi)\}$$

and evaluates the marginal density function

$$p_{\mathbf{Z}}(z; \xi) = \int p_{\mathbf{Z},\mathbf{X}}(z, x; \xi) dx.$$

As usual, the algorithm actually works in terms of the negative logs of the various probability densities.

In each of our experiments, we used the cited algorithm to define an objective function  $f_z(\xi) = -\log(p_{\mathbf{Z}}(z; \xi))$ , that depends on the measurement sequence  $z$ , and then used Powell's conjugate direction minimization algorithm (Chapter 10, Section 5 of [16]) to find the optimal parameter values for each measurement sequence  $z$

$$\xi_z^* = \operatorname{argmin}_{\xi} \{f_z(\xi)\}.$$