Wi-Fi Location Determination for Semantic Locations

Keith J. Kelley
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/hilltopreview

Part of the Computer Sciences Commons

Recommended Citation
Available at: https://scholarworks.wmich.edu/hilltopreview/vol7/iss1/9
Wi-Fi Determination Location for Semantic Locations

By Keith J. Kelley
Computer Science Department
keith.kelley@wmich.edu

1 Introduction

The prevalence of smartphones has increased demand for indoor positioning where GPS cannot reach. Because Wi-Fi has become ubiquitous, it is a natural technology to turn to for indoor location determination. Wi-Fi location determination, also known as Wi-Fi localization or Wi-Fi location estimation refers to methods of translating observed Wi-Fi signal strengths into locations. A “radio map”, consisting of sets of Received Signal Strength Indicators (RSSI) is stored as a “fingerprint” to be compared later to a signal scan to recognize the location of the device doing the scanning.

In Wi-Fi location determination literature, little attention is paid to locations that do not have numeric, geometric coordinates, though many users prefer the convenience of non-coordinate locations (consider the ease of giving a street address as opposed to giving latitude and longitude). It is not often easy to tell from the title or abstract of a Wi-Fi location determination article whether or not it has applicability to semantic locations such as room-level names.

The concept of a semantic location was defined in a paper from HP Labs, (Pradhan 2000), specifically to address an important deficiency—physical locations have little context information for mobile web.services. Pradhan defined a semantic location as a URI with context information. Semantic location have since been redefined through use as any symbolic location that can be expressed in natural language as stated in (Darabi et al. 2007), or that have particular meaning to users or applications (Roth 2005), as opposed to the use of geometric or physical coordinates (Roth 2005). An article is applicable to semantic locations if it is applicable to symbolic locations; semantic locations are symbolic locations that possess a single coordinate that is a symbol that can be understood by another person or piece of software.

Semantic locations, then, contrast with other types of locations, such as relative (defined in relation to a known reference point), absolute (defined by position on a coordinate grid shared with other locations) and physical (defined by geospatial coordinates—latitude and longitude) (Darabi et al. 2007). Other articles use physical as a synonym for geometric locations. The four types of locations can be divided into two classes of coordinate systems, symbolic and geometric, as in (Becker and Dürr 2005). Symbolic or semantic locations have symbolic coordinates, whereas absolute, relative, and physical locations are all defined by a set of geometric (x,y) coordinates. Geometric coordinates implicitly define distance between two objects, while symbolic coordinates do not. (Hightower and Borriello 2002) likewise contrasts physical position and symbolic location, as well as absolute versus relative. Given that three of the four types of locations defined in Wi-Fi localization articles have geometric coordinates and the fourth has symbolic coordinates, it is then logical that location systems built around semantic locations would be different.

1.1 Needs Met By Semantic Locations

Location is well known to be an important source of context for context-aware applications and is completely vital for the location-based services subset. (Roth 2005) delineates “The Role of semantic locations for Mobile Information access” as distinct from other location types or uses. Semantic locations are needed because in lieu of geometric
information, symbolic information is vital to location-based services. Systems that use semantic locations directly are needed:

- geometric x,y coordinates aren’t always available to users or builders of location-aware systems
- symbolic locations often represent areas, whereas physical locations represent points (Roth 2005)
- not every location is the same size or shape
- not every location is stationary (for example “My Car,” “Flight 405”)
- semantic coordinates are easier to understand and more useful than physical coordinates “Semantic Location,” (Pradhan 2000) (Roth 2005)
- in many apps, symbolic locations are closer to what the user wants (GPS coordinates have to be translated so it is known they represent a coffee shop)
- in many instances, it is more difficult to get a physical location than a semantic one
- it makes no sense to waste cycles finding physical location information if your end goal is a semantic location
- rooms are seen by many users as being more important than geometric coordinates while indoors (Yang, Lu, Jensen 2009)
- geometric movement models suitable for outdoors are not suitable for indoors (Yang, Lu, Jensen 2009)
- indoor positioning is different and more difficult than outdoor positioning

The goal of a semantic location is different than the goal of a physical coordinate. The former’s goal is to communicate symbolic information to a user or application and the latter’s goal is to return a position in space that can be used for mapping or other distance-oriented applications.

1.2 Complications of Using Semantic Locations

Semantic locations simplify some aspects of system design, while complicating others. Many positioning methods require geometric information at certain reference points and cannot be used for symbolic locations. (Pradhan 2000) notes that symbolic locations can be ambiguous, difficult to sense, and too coarse for many applications.

1.3 Using Geometric Locations to Get Semantic Locations

Much effort has gone into turning physical locations from services like GPS into semantic locations for location-based services. (Cao, Cong, Jensen 2010) and (Juhong Liu, Wolfson, Huabei Yin 2006) focus on using physical coordinates to obtain semantic locations for use in mobile applications. Indeed, many commercial applications use GPS antenna or network-provided latitude and longitude to determine if a user is at a specific semantic location. In many cases this method will work, but in some cases, it will only waste battery time, and in others (such as “My Car”) may not even be viable.

Simultaneously, much effort has gone into finding physical locations for indoor locations, and as most GPS is unavailable indoors; no small amount of that effort going into Wi-Fi location determination. As physical location is often a means to get semantic location, it makes sense to skip the extra step for applications that can use semantic locations.

1.4 History of Semantic Locations in Wi-Fi Location Determination

The earliest Wi-Fi location determination systems did not consider semantic locations. The first algorithm, NNSS, supported them implicitly, though only as a mechanism to get to geometric coordinate information, and the second algorithm, k-NN (referred to as NNSS-AVG), implicitly excluded their use, requiring geometric coordinates (Bahl and Padmanabhan 1999). Prior to 1999 localization was largely a robotics domain, usually requiring geometric
coordinates and more recently, focus has been on acquiring physical coordinates because semantic locations are adequately provided by proximity to a beacon.

1.5 Uses of Semantic Locations
Semantic locations are usually associated with smart homes and the “internet of things.” People want certain services associated with locations, for example to change the settings on their phone when they get to “the office” or to turn on the TV when they get to “the living room” or to turn on the lights when they reach “home.”

2 Wi-Fi Location Determination Methods
A few articles on Wi-Fi location determination explicitly mention symbolic or semantic locations, such as (Mantoro et al.) and (Mantoro and Johnson 2005), but often one must determine the applicability based on the algorithm.

There are three main types of Wi-Fi location determination methods: proximity, triangulation, and fingerprinting (or scene analysis).

2.1 Proximity
The proximity method uses a mobile device is within range of a fixed antenna for positioning. Proximity, by nature, gives symbolic coordinates (near antenna X), but if the antenna locations are known, they are often mapped to physical coordinates. Of course, the accuracy is no better than the radius of the antenna’s range. While that makes sense for short-range systems like the specialized Active Badge or Bluetooth, Wi-Fi location determination methods do not often use the proximity method because of the relatively long range of Wi-Fi, the cost, and the interference problems inherent in having too many wireless access points in the same physical space. So, while proximity methods inherently give symbolic coordinates, they have limited use when combined with Wi-Fi.

“Indoor Positioning with a WLAN Access Point List on a Mobile Device” describes a proximity-based method where each MAC address is mapped to a floor or wing of a building. No reference to symbolic or semantic locations is made.

A recent commercial standard is the Bluetooth LE iBeacon from Apple.

2.2 Triangulation
Triangulation is use the use of the angles and the length of the sides of triangles with basic trigonometry to determine a point based on 2 or more sources. Triangulation methods are by definition geometric and require distance coordinates to work. Lateration methods only require distances, but angulation methods also require angles to work. Since triangulation methods require geometric coordinates, they cannot be used directly with semantic locations.

2.3 Fingerprinting
Wi-Fi location fingerprinting takes a “fingerprint” of all visible antennas and their strengths as an identifier of how the radio signals should look at that location. Fingerprinting methods, a type of scene analysis, can sometimes support semantic locations and the determination is on a case-by-case basis. If a method requires manipulation of the geometric coordinates of a location, it cannot be applied to semantic locations, though if it can be modified to remove that requirement, it can be applied.

Fingerprinting can be either deterministic or probabilistic. Of the three main deterministic methods, only one is usable for symbolic locations because the others require geometric coordinates.
2.3.1 Deterministic Methods

2.3.1.1 Nearest Neighbor

Nearest Neighbor in Signal Space (NNSS), more commonly referred to as NN, is a symbolic location determination algorithm, as no averaging of geometric coordinates occurs. The signal strengths observed by the mobile device are simply measured against the signal strengths previously recorded for the same antennas, and the closest antenna in distance is chosen as the nearest neighbor. It was first introduced in (Bahl and Padmanabhan 1999) and works for semantic locations, though no special attention was given to the fact that NNSS does not require geometric coordinates. NNSS is based on the more commonly known Nearest Neighbor Search (NNS), also sometimes referred to as NN.

(Kaemarungsi and Krishnamurthy 2004a) presents a mathematical model of NN on a grid of locations to the end of designing such a system. This has direct applicability to symbolic space, as the primary criterion of success is probability of correct location estimation, but is an idealized case. The localization system designer is also the wireless network designer and it is only applicable to nearest neighbor, which is not the best available algorithm. While it might be rare that a location determination system would conform to a grid, the article is applicable to semantic locations on such a grid and some of article might be more generally applied.

One enhancement to NNSS is in “Wireless Indoor Positioning System with Enhanced Nearest Neighbors in Signal Space Algorithm,” which adds the additional criteria of a signal strength threshold and the number of antennas that may deviate beyond it.

Not every modification of NNSS can be applied to semantic locations, however, as some, such as kNN, add a requirement that geometric coordinates be used in the computation.

2.3.1.2 kNN

(Bahl and Padmanabhan 1999) contains a description of the first version of kNN for Wi-Fi fingerprinting (called NNSS-AVG) which doesn’t apply to semantic locations, as it averages the coordinates of the k=3 nearest neighbors in signal space without any special attention paid to the new requirement that the reference point’s physical coordinates be known.

(Yang, Lu, Jensen 2010) uses a kNN search but in the original sense (kNN outside of Wi-Fi localization is an old algorithm with much more general application in pattern recognition), but isn’t actually a Wi-Fi method at all; they assume the use of a symbolic proximity detection method such as Bluetooth or IR.

In pattern recognition, kNN is the more general form of NN and it is such a popular method for Wi-Fi location determination that even methods based on NNSS often say they are based on kNN with k=1. (Lee et al. 2010) is an example of an article that states that it uses kNN, implying its inapplicability to semantic locations, but upon closer reading, is actually based on NNSS. Redpin is another such system (http://www.redpin.org)

2.3.1.3 WKNN

WKNN, also referred to as KWNN, is the Weighted k-Nearest Neighbors algorithm, which averages the coordinates of the k nearest neighbors together, adopting the distances in signal space as weights. Since semantic locations have no geometric coordinates to weight together, methods that are a variant of KWNN or WKNN do not apply to semantic locations.

There are conceivably many more deterministic methods than there are programmers, but a simple rule of thumb applies: if the algorithm requires geometric coordinates to work, it cannot usually be used for semantic locations as the output of the algorithm is typically a coordinate and not a location name. Also, some modifications to kNN will not reap any benefits over NN if K is set to 1, so when considering a kNN variant, it is important to be sure of the impact of keeping that as a fixed constant.
2.3.2 Probabilistic Methods

Probabilistic location determination methods take the measured set of signal strengths of nearby antennas and compute the likelihood of each location in the fingerprint database being the current nearest location. Any probabilistic approach can be used, as long as only the most probable location can be computed by itself. Normally, in the case of relative, absolute, or physical locations, a probabilistic method interpolates the position coordinates by using the weighted average of the coordinates of all calibration points times their likelihoods (Darabi et al. 2007).

(Castro et al. 2001) is an example of a probabilistic system that returns semantic locations in the form of a most likely room name based on the time of day that the matching fingerprint was taken.

(Yang, Lu, Jensen 2010) is an example of a probabilistic method influenced by the kNN and named after it, while not even being deterministic.

(Ladd et al. 2004) uses a probabilistic Bayesian technique, which returns a trained position or state.

(Youssef et al. 2002) is applicable to symbolic locations, but makes no mention of the concept, nor really distinguishes in any way between types of locations. Two methods are described, a joint clustering technique and something called iterative trilateration, but with distances based in signal space, not physical space. The Joint Clustering Technique is further described in (Youssef, Agrawala, Shankar 2003)

(Roos et al. 2002) does not mention semantic or symbolic locations, but refers to labeling locations. The Nearest Neighbor method is compared to the histogram method and the Gaussian kernel method.

2.4 Multiple Methods

Many articles compare different methods, such as (Lee et al. 2010) which compares a probabilistic method to kNN with k=1.

(Navarro, Peuker, Quan 2010) is a senior project that tracks children on a playground and uses both nearest neighbor, a deterministic method and Markov Localization, a probabilistic method, to return a most likely location.

(Widyawan and Pesch 2007) uses the same one slope model and multi-wall models used in (Narzullaev, Park, Jung 2008). In the positioning phase the authors use NN as well as a particle filter method.

(Saha et al. 2003) considers location determination as a classification problem and evaluates a nearest neighbor classifier, a neural network, and a histogram matching classifier. Errors are measured in distance, but results are given by location.

(Correa et al. 2008) provides “Room-Level Wi-Fi Location Tracking,” and compares deterministic and probabilistic methods, concluding that it is a good compromise to store full histograms rather than averages or full result sets.

(Badawy and Hasan 2007) uses three methods: Nearest Neighbor, a neural network, and a decision tree. While not stated, the locations in question are symbolic, as evident by the fact that accuracy is measured in percentage instead of meters.

(Lin et al. 2009) uses a naive Bayes classifier, an SVM, kNN with k=1, Redpin, and introduces WASP, all of which are methods of using Wi-Fi to determine semantic location. Rather than referring to symbolic or semantic locations, however, the article refers to rooms. It is clear that the authors have made note of the distinction, though, as they point that a benefit of fingerprinting is that the designers of the system do not need to know the physical locations of all of the wireless APs.
The Hilltop Review, Winter 2014

(Smailagic and Kogan 2002) uses both CMU-TMI (triangulation) and CMU-PM (fingerprinting), which is originally described in (User-Centered Interdisciplinary Design of Wearable Computers).

2.5 Other Methods

(Schloter and Aghajan 2006) implements a Support Vector Machine (SVM) for symbolic positioning in order to bypass the added requirements of a system that uses physical coordinates, which is acknowledged as an intermediary step along the way to symbolic locations for most uses.

(Yim 2008) describes a decision tree method the author compares to 1-NN, Bayesian, and other decision tree-based methods. While not explicitly stated, it is certainly more suited to symbolic locations than geometric.

3 Applying the Literature

In general, appropriate methods can choose a nearest neighbor without using physical or geometric coordinates or the geometric coordinates of any reference points. This includes any deterministic Wi-Fi fingerprinting method if k=1 (assuming a variant of kNN), as well as any probabilistic Wi-Fi fingerprinting method (without the interpolation or averaging step), returning just the single location with the greatest likelihood. Any algorithm that requires and returns geometric coordinates cannot be applied to semantic locations. This includes any WkNN variant, any triangulation method, and most kNN variants.

As we have seen, not every system is applicable to semantic locations. To date, most articles on wireless location determination have made little distinction between symbolic and geometric locations. The reader must take into account the considerations above for each method mentioned in a paper or a survey before knowing if it is applicable.

3.1 Taxonomy

Articles written about location determination systems for symbolic locations are often unintentionally obscured by their authors due to the lack of a specific vocabulary.

(Kjærgaard 2007) is an example of a proposed taxonomy that does not classify location determination systems by location type. It specifies area versus point or path as one of three spatial properties of the collection method. This fails to bring to the forefront the fundamental importance the distinction has on the rest of the system.

(Zeimpekis, Giaglis, Lekakos 2002) comes from a business perspective, where semantic locations are highly relevant, but also does not distinguish between geometric and semantic locations.

3.1.1 Accuracy

Accuracy and precision are often defined differently in location determination. Location accuracy is usually defined as error distance between the real location and the estimated one. Location precision is usually defined as percentage of correct estimations at a specific error distance. Location accuracy as typically defined for location determination has little or no meaning in symbolic location systems. Accuracy in articles specifically about systems for symbolic locations is just as often defined as percentage of correct location identifications, which is normally called location precision.

Distance, by any name, is an almost-ubiquitous error measure, even used by researchers whose systems have no concept of distance.

(Elnahrawy, Li, Martin 2005) actually uses different criteria for area matching, including its own definitions of accuracy and precision. The researchers make no reference to semantic or symbolic locations, referring instead to areas, tiles, or rooms. Rather than physical or
geometric coordinates, they refer to points. They developed three algorithms for area and compared to kNN as a representative of point-based algorithms.

### 3.2 Articles That Implement Methods For or Make Use of Semantic Locations

The following articles compare two or more different localization methods.

- (Lee et al. 2010) uses kNN with $k=1$ and a probabilistic method.
- (Ching et al. 2010) describes a system in a university environment with room-to-room accuracy, without mentioning semantic or symbolic locations. The authors used an unspecified deterministic algorithm that may be NNSS.
- (Anderson and Muller 2005) uses symbolic locations by automatically creating zones, as well as using fingerprinting and a Bayesian Network.
- (Hansen et al. 2009) describes how to configure semantic location determination in Streamspin, a mobile services platform for Windows Mobile devices.
- (Kelley, Kaugars, Garrison 2011) makes use of any location algorithm for semantic locations.
- (Yuen, Balasubramaniam, Din 2010) contains a method for improving kNN that could be applied to NN. The method deals with filtering out fluctuation in Received Signal Strength (RSS). RSS is typically indicated by a Received Signal Strength Indicator, an integer between -255 and 0 in theory, but commonly between -30 and -100 which either is or is proportional to the dBm, depending on the implementation.
- (Sabbour 2007) is a 2007 Bachelor’s Thesis that implements two deterministic symbolic location algorithms for fingerprinting: Bahl’s Nearest Neighbor with Euclidean distance, and a range-based algorithm, which is based on each antenna’s RSS being within a specific range. The range-based algorithm chooses the room for which the most training samples match. NNSS was found to be accurate 10% more often.
- (Jr. 2003) proposes a reference database of symbolic locations using doorways as reference points.

### 3.3 Surveys and Comparisons

- (Darabi et al. 2007) and (Gezici 2008) both cover areas involving wireless position estimation, but not specifically semantic or symbolic locations.
- (Gu, Lo, Niemegeers 2009) includes symbolic systems using a variety of technologies, such as RADAR (Wi-Fi), Active Badge (IR) (symbolic), and Cricket (Ultrasound) (semantic).
- (Seco et al. 2009) describes kNN and Probability Density Function methods of fingerprinting, and explains that machine learning methods are better for symbolic locations than for physical (geometric) coordinates.

A number of comparisons of indoor positioning methods do not mention symbolic or semantic locations by any term, but do describe methods that can be applied to them. They also use mean or median distance error as their criteria for accuracy comparison, which has little applicability to locations with non-geometric coordinates (Lin and Lin 2005) (Wallbaum and Diepolder 2005) (Li et al. 2006) (Leppäkoski et al. 2010) (Honkvirta et al. 2009).

One notable exception is (Darabi et al. 2007), which describes the four location types, including symbolic. The authors still maintain the convention of measuring error in meters as a measure of accuracy, though precision, defined therein as percent correct within 0 meters, is also included.

That only one of these acknowledges the research into symbolic locations shows how little attention is paid to this area.

### 3.4 Articles That Apply Somewhat to Semantic Location Determination
(Swangmuang and Krishnamurthy 2008a) has some methods which can be applied to
fingerprints for locations of any type.

(Bolliger et al. 2009) is applicable to improving accuracy of semantic location
determination systems through better training of the radio map without additional effort. A
symbolic location in this article is referred to as a label (for a location).

(Swangmuang and Krishnamurthy 2008b) focuses on cleaning up the fingerprint database
(radio map), which could be applied positively to a method like NNSS, but could actually
have a negative impact on probabilistic methods.

(Cook et al. 2009) applies just as well to symbolic or physical coordinates.

3.5 Articles That Apply to All Location Types

(de Moraes and Nunes 2006) use a probabilistic method of fingerprinting. Versions of
this method have been used in Horus and other systems. The probability of the most likely
location is calculated using Bayes’ rule. The goal of this project is not a symbolic location, so
it goes on to use one of two further methods to calculate physical position.

(Prasithsangaree, Krishnamurthy, Chrysanthis 2002) looks at ways of improving both
symbolic and geometric locations, and describes them without ever using specific
classification terms. (Essentially, NN and kNN are described). Improving NN by using
weights of the first number of signal samples, then standard deviation was tested. They found
no improvement with these weights.

(Leppakoski, Tikkinen, Takala 2010) examines the effect of histogram bin size in
probability-based fingerprinting methods. There is no reason that the results wouldn’t apply to
symbolic locations as well, though the impact on accuracy would be different.

(Tsui, Chuang, Chu 2009) is something that could be applied equally to improving
symbolic or geometric location determination methods.

(Kjaergaard and Munk 2008) is a solution to the problem of hardware RSS perception,
which could be applied to symbolic, as well as geometric locations.

Filters such as those presented in (Suárez, Elbatsh, Macías 2010) are applicable to either
kind of location.

3.6 Articles about RSSI

Most articles that focus on overcoming some problems of RSSI and Wi-Fi Location
Determination apply equally well to symbolic and physical locations.

(Kaemarungsi and Krishnamurthy 2004b) is still applicable to symbolic of course.

(Chan, Baciu, Mak 2010) defines accuracy as percentage correct, so is more applicable to
symbolic than geometric.

(Fang et al. 2010) is not all that useful for fingerprinting for symbolic locations, though
there are commonalities.

(Tsui, Chuang, Chu 2009) could be used for symbolic locations.

3.7 Open Source Systems

Redpin (http://www.redpin.org) does provide symbolic identifiers, such as room name or
number, and is said to be based on kNN where k=1. In other words, Redpin is based on
NNSS.

Herecast (http://sourceforge.net/projects/herecast/) is an open source symbolic location
system on Pocket PC hosted for providing location-based services (Paciga, 2005).

PlaceLab (http://sourceforge.net/projects/placelab/) is another open source system
developed by Intel Research that ran on mobile devices including PDAs, phones and other
devices that was meant to be local to the device and for large-scale environments (LaMarca,
2005).
3.8 Commercial Systems

A large number of commercial systems have been created that make use of or deliver semantic locations, including Foursquare, Facebook Places, and Twitter Places. Most of these systems use physical position and define the symbolic locations on top of those.

(Wexler 2006) looks at a handful of commercial systems.

3.9 Recent Developments

Especially in the retail space, in the last couple of years Apple's iBeacon has largely taken the place of other technology for location based services, however, research has not stopped completely.

(Baniukevic 2013) use both Bluetooth and Wi-Fi for location estimation, with the Bluetooth proximity representing a symbolic location at times and the Wi-Fi allowing the representation of a position on a grid.

Ariel (Jiang 2012) is a system for room localization that automatically learns fingerprints based on Wi-Fi and motion.

(Trawiński 2013) used a fuzzy rule-based multiclassification system to do Wi-Fi fingerprinting without knowledge of the environment nor additional infrastructure.

(Husen 2014) used Wi-Fi RSSI to establish both position and orientation of the users.

4 Conclusion

To this day, commercial apps for consumers like Foursquare are more reliable for building level locations, but benefit when they can provide services at the room level or below, and GPS signals will never reach inside many buildings. Even though many are interested in the location based services provided by having semantic locations indoors research into semantic locations seems to have fallen off in the last couple of years, as the iBeacon and competing or compatible systems have been released. An iBeacon allows you to identify that you are in the proximity of a particular place. A place with an Bluetooth LE iBeacon or equivalent. A building or campus with saturated WiFi and a good method for using that WiFi to determine location does not require the added infrastructure cost of a beacon at every location, but accomplishes the same thing with nothing more than the WiFi antennas already in place. In such an environment, or in environments where theft or vandalism of beacons may be an issue, iBeacons may have been obsolete before they were announced. However, with research in this area largely having fallen off it seems iBeacon will become the standard for identifying semantic locations for location based services as it is cheap enough and easy enough for most anyone to implement.

References


Mantoro T and Johnson C. 2005. Nk-nearest neighbor algorithm for estimation of symbolic user location in pervasive computing environments. World of Wireless Mobile and
Multimedia Networks, WOWMOM '05 Proceedings of the Sixth IEEE International Symposium on. IEEE. p472-474


Sabbour AA. 2007. WiGuide: Indoor system for LBS. Faculty of Media Engineering and Technology, German University in Cairo.


