ABSTRACT
A practical WiFi-based positioning system has to be adaptable to the variations of indoor environmental dynamic factors. In this work, we propose a novel Wi-Fi indoor positioning and tracking framework which employs the spatial analysis and image processing techniques. The Wi-Fi surfaces can be dynamically constructed and updated and thus help to address the challenges of signal spatial heterogeneity and environmental variations. A mobile app for indoor positioning application has been developed as a proof of concept. Based on the experiments we conducted at the Esri campus, this method can achieve about 2-meter positioning accuracy. The proposed methodology and theoretical framework can guide engineers to implement cost-effective indoor positioning infrastructure, and thus offer insights on future smart campus applications. The introduced spatial analysis and geoprocessing workflow may also bring the attention of GIScientists to make more efforts to conquer the indoor positioning and tracking challenges.

CCS Concepts
- Information systems → Mobile information processing systems; Spatial-temporal systems; Location based services; Geographic information systems; Global positioning systems;

Keywords
Indoor positioning; Spatial analysis; WiFi fingerprinting

1. INTRODUCTION
In the era of mobile age, location-based systems and services become more and more popular in people’s daily life, such as searching nearby points of interest (POI), way finding and navigation. There has been growing interest among researchers in studying human mobility patterns based on the data collected from location-awareness devices, e.g., GPS-enabled devices [45, 39], cellular phones [13, 18], and Bluetooth sensors [29, 33]. Positioning is the key component to support smart mobility studies, trajectory data mining and travel related applications [43, 36, 40]. It is well known that GPS-enabled devices work well for positioning in outdoor environments but not in indoor environments because of the signal obstructions and attenuation[26]. However, as reported in a national human activity pattern survey (NHAPS), people spend an average of 87% of their time in enclosed buildings and about 6% of their time in enclosed vehicles [19]. There is a high-demand for indoor positioning technologies. In [44] the researcher systematically compare three dominant positioning technologies: Assisted-GPS, Wi-Fi, and Cellular positioning. Their pros and cons are discussed in terms of coverage, accuracy and reliability. It reports that Assisted-GPS obtains an average median error of 8m outdoors while Wi-Fi positioning only gets 74m of that and cellular positioning has about 600m median error in average and is least accurate. However, high-resolution GPS or Assisted-GPS positioning chipsets don’t work well in indoor environments due to limited satellite visibility; and thus a variety of indoor positioning technologies and systems have been designed and developed to increase the indoor positioning accuracy.

The widely use of Wi-Fi access points for Internet connection in hotels, offices, coffee shops, airports and many other fixed places makes Wi-Fi become an attractive technology for the positioning purpose. Location positioning systems using wireless area local network (WLAN) infrastructure are considered as cost-effective and practical solutions for indoor location estimation and tracking [7]. There is almost no extra hardware or other infrastructure investment, which is different from other indoor positioning technologies such as low-energy Bluetooth sensor networks or radio-frequency identification (RFID) systems. Conventional methods for indoor positioning are based on time of arrival, time difference of arrival and angle of arrival of radio signals transmitted by mobile stations [30]. Another type of WLAN positioning method unitizes the fingerprinting idea to determine a receiver’s location via clustering and probabilistic distributions [42]. The accuracy of localization is dependent on the separation distance between two-adjacent Wi-Fi reference points and the transmission range of these reference points [5]. One challenge that WLAN indoor positioning techniques face is the spatial heterogeneity of Wi-Fi signal surfaces caused by the signal path loss because of absorption, reflection and refraction by the surrounding environment. Thus the physical radiation models need complex parameter estimation and empirical fitting. Spatial analysis
techniques can be employed in dealing with geometric, topological, and geographic properties which can help to capture better indoor environment and to construct more accurate Wi-Fi surfaces for indoor positioning and tracking purpose.

In this research, we propose a novel geoprocessing-based framework and develop a mobile application for real-time indoor positioning and tracking using Wi-Fi access points. The goal of this study is to explore how accurate Wi-Fi indoor positioning with the support of spatial analysis can achieve. The following of this paper is organized as follows. In Section 2, we review some related work and discuss the state of the art in indoor positioning technologies. Our proposed theoretical framework towards indoor positioning and tracking, and detailed methods are introduced in Section 3. The system implementation and experiments are described in Section 4, which is followed by discussions in Section 5. We conclude and give a vision for future work in Section 6.

2. LITERATURE REVIEW

2.1 Indoor Positioning Technologies

In general, indoor positioning technologies can be classified into two broad categories: radio-frequency-based (RF) and non-radio-frequency-based (NRF) technologies. The RF group includes but not limited to WLAN, Bluetooth, and RFID systems, while the NRF group contains ultrasound, magnetic fields, and vision-based systems. Researchers have made great efforts in the field of indoor positioning using these sensors and technologies [17, 11, 23, 15, 21, 24, 20]. The spatial coverage area and positioning accuracy of those different technologies have been reviewed by [26]. Here, we only briefly discuss the RF technologies that are most popular in the current market share and face several challenging issues to investigate.

Wi-Fi: The Institute of Electrical and Electronics Engineers (IEEE) 802.11 work group [14] documents the standard use of Wi-Fi technology to enable wireless network connections in five distinct frequency ranges: 2.4 GHz, 3.6 GHz, 4.9 GHz, 5 GHz, and 5.9 GHz bands. The wireless networks are widely implemented in many types of indoor buildings in which wireless access points are usually fixed at certain positions. Those access points allow wireless devices (e.g., mobile phones, laptops and tablets) to connect to a wired network using Wi-Fi technology. And the relative distance between wireless devices and access points can be roughly estimated based on Wi-Fi signal strength using signal propagation models [28]. It is also well known that the accuracy of indoor position estimation based on Wi-Fi signal strength is affected by many environmental and behaviour factors, such as walls, doors, settings of access points, orientation of human body, etc. [11, 37]. In practical applications, a good approximation of heterogeneous environmental signal surfaces could help to improve the indoor positioning accuracy. Spatial regression which is a widely used spatial-analysis method in finding spatial patterns and modeling trend surface [8] could potentially be a good candidate.

Bluetooth: is designed for low power consumption and allows multiple electronic devices to communicate with each other without cables by using the same 2.4 GHz radio-frequency band as Wi-Fi. The distance range within which Bluetooth positioning can work is about 10 meters. In the beaconing mode, Bluetooth permitted messages can be used to detect the physical proximity between two devices [10].

In an indoor environment equipped with equal to or large than three Bluetooth low energy (BLE) beacons, the location of a target mobile device with Bluetooth can be determined using classic positioning approaches (see Section 2.2 in detail). In this way, location-dependent triggers, notifications and tracking activities can be enabled by employing multiple BLE beacons. There are several popular BLE beacon-positioning protocols and technology available online, including Apple iBeacon1, Google Eddystone2 and Qualcomm Gimbal3, which guide developers to implement up-to-date indoor positioning and tracking applications.

Radio-frequency Identification (RFID): is a general term used for a system that communicates using radio waves between a reader and an electronic tag attached to an object. Comparing with Bluetooth technology, RFID systems usually comprise of readers and tags that store relatively limited information about the object such as location and attribute information. Those tags can be activated and send out stored information if they receive the signal from RFID readers within certain distance thresholds, which can be used to estimate the reader’s location and to show relevant information. Currently RFID positioning and tracking systems are widely used for asset tracking, shipments tracking in supply chains, and object positioning in retailing places and shopping malls.

Because of sensor diversity and positioning challenges in various indoor environments, there is also an increasing trend towards combining and integrating different sensor networks to get a better spatial coverage and position accuracy than using single data source. In [9] researchers integrate Wi-Fi and inertial navigation systems to get performance close to meter by fusing pedestrian dead reckoning and Wi-Fi signal strength measurements.

2.2 Indoor Positioning Methods

The following methods usually work for both outdoor and indoor environments, but we emphasize the indoor case here. Note that those methods can be applied in most radio-frequency-based sensors, such as Wi-Fi and Bluetooth.

2.2.1 Geometric Approaches

In geometry, trilateration is a method to determine the target location of a point by measurement of distances to three points at known locations using the geometry of circles, triangles, or spheres [6]. This method has been widely used not only in positioning systems [16], but also in robot localization, computer graphics, aeronautics, and so on [34]. For the positioning purpose, if already knowing the coordinate (lat,lon) information of three fixed access points, it needs to convert the latitude and longitude of these locations from the Earth reference system to axis values (x,y,z) in the Cartesian coordinate system as follows:

\[ x = R \times \cos(lat) \times \cos(lon) \]
\[ y = R \times \cos(lat) \times \sin(lon) \]
\[ z = R \times \sin(lat) \]

Where \( R \) is the approximate radius of earth (e.g., 6371Km). Then we can try to find the solutions to trilateration equations to approximate the location of the target

1https://developer.apple.com/ibeacon/
2https://developers.google.com/beacons/
3http://www.gimbal.com/
or less than the maximum number of available Wi-Fi access points at known reference point (RP) locations [17, 22, 27, 12]. In its simplest form, it can be expressed mathematically as follows:

\[
\text{argmin}_{p \in \{1, 2, n\}} \sum_{i=1}^{m} \left[ SS_{RE}(i, p) - SS_{ME}(i) \right]^2
\]

where \( SS_{RE}(i, p) \) represents the received signal-strength value of access point \( i \) at a known reference point location \( p \) on the radiomap, and \( SS_{ME}(i) \) is the measured signal-strength value of access point \( i \) at the current unknown location. The location \( p \) that has the minimum root-squared-differences of SS values for all available \( m \) access points between reference points and the target location is considered as the most probable estimated location.

In addition, machine-learning-based fingerprinting techniques have also been studied for improving the quality of location estimation in complex real environment, including Bayesian modeling, k-nearest-neighbor estimation, support vector machine, neural networks and so on [2, 31, 42, 4].

3. METHODOLOGIES

3.1 Study Design

Our theoretical framework of indoor positioning and tracking consists of three steps as shown in Figure 2.

Step 1: Field sampling is a process of collecting available indoor Wi-Fi access point set \( \{AP_1, AP_2, AP_m\} \) and the corresponding received signal strength indicator (RSSI) at various locations \( \{Loc_1, Loc_2, ..., Loc_n\} \) with different distances to each access points. The output of this step is a two-dimensional matrix \( \{RSSI(AP_i, Loc_j)\} \) which can be represented as:

\[
\begin{bmatrix}
RSSI(AP_1, Loc_1) & RSSI(AP_1, Loc_2) & \cdots & RSSI(AP_1, Loc_n) \\
RSSI(AP_2, Loc_1) & RSSI(AP_2, Loc_2) & \cdots & RSSI(AP_2, Loc_n) \\
\vdots & \vdots & \ddots & \vdots \\
RSSI(AP_m, Loc_1) & RSSI(AP_m, Loc_2) & \cdots & RSSI(AP_m, Loc_n)
\end{bmatrix}
\]

Note that the matrix allows for ‘null’ value element if there is no signal for a given Wi-Fi access point at a certain location.

Step 2: The fixed location of each Wi-Fi access point \( Loc(AP_i) \) can be estimated based on a sample of high quality signals (i.e., RSSI value ≥ -50 decibels). The geometric centroid of these clustered high quality signal locations is stored as \( Loc(AP_i) \). Meanwhile, as shown in the geoprocessing workflow (Figure 3), the signal surface of each Wi-Fi access point \( Surf(AP_i) \) can be built via spatial interpolation (e.g., inverse distance weighted (IDW) method or Kriging techniques) of collected sampling points from Step 1. The AP locations and raster surfaces of these Wi-Fi access points generated from this geoprocessing step are stored on the GIS Cloud server where data can be published as

object by referring to three points with known locations [16] (See Figure 1). The equations are nonlinear and it is not so easy to obtain an exact solution. Several iterative arithmetic methods have been proposed to find efficient solutions for trilateration-based localization [25, 41].

It is obvious that the distance between a wireless device and an access point is the key for Wi-Fi positioning using the trilateration method. The received signal strength doesn’t directly lead to an estimated distance to an access point. In general, it does follow a trend that the signal strength decreases with an increase of distance as is expected, but it is not a simple linear path loss model. The log-distance path loss model is one of the most simplistic radio-propagation models [28, 32] that predict the received signal strength at a certain distance inside a building or densely populated areas. However, in many practical situations, many factors underlying radio propagation can contribute to the reflection, refraction, absorption and scattering of signals. It is difficult to predict the received signal strength with a simplistic log-distance pass-loss model. Rather than creating a new pass-loss model, some researchers try to re-parameterize the classic log-distance model. For example, researchers found that at closer ranges (e.g., smaller than 5 meters) the exponent factor of propagation model would take a higher value, and thus a dual-distance model has been proposed to adjust the propagation model at larger distances for achieving a better positioning accuracy [3].

2.2.2 Fingerprinting Approaches

Although most Wi-Fi access points are located at the fixed locations of buildings, it is very hard to get their accurate coordinates and detailed digital floor maps because of privacy concerns. Fingerprinting approach is a popular wireless-network positioning technology in metropolitan areas since it doesn’t need the exact position information of Wi-Fi access points. Instead, it needs to construct an offline database that contains the signal strength distribution of Wi-Fi access points at known locations, i.e., the “radiomap”. For a given location, it receives varying signal values from equal to or less than the maximum number of available Wi-Fi access points. The set of access points and their signal strengths are distinctive and present a “fingerprint” to a unique location on the “radiomap”. Thus, we can employ searching and comparison processes in the online positioning phase to find the most probable location.

Several fingerprinting matching and calculation methods have been developed to estimate the location of a user based on received Wi-Fi signal strength (SS) values at known reference location, i.e., the “radiomap”. For a set of access points that contains the signal strength distribution of Wi-Fi access points. Instead, it needs to construct an offline database that doesn’t need the exact position information of Wi-Fi access points. Rather than creating a new fingerprinting model, some researchers try to re-parameterize the classic log-distance pass-loss model. For example, researchers found that at closer ranges (e.g., smaller than 5 meters) the exponent factor of propagation model would take a higher value, and thus a dual-distance model has been proposed to adjust the propagation model at larger distances for achieving a better positioning accuracy [3].

\[
\begin{align*}
& \text{argmin}_{p \in \{1, 2, n\}} \sum_{i=1}^{m} (SS_{RE}(i, p) - SS_{ME}(i))^2 \\
& \text{where } SS_{RE}(i, p) \text{ represents the received signal-strength value of access point } i \text{ at a known reference point location } p \text{ on the radiomap, and } SS_{ME}(i) \text{ is the measured signal-strength value of access point } i \text{ at the current unknown location. The location } p \text{ that has the minimum root-squared-differences of SS values for all available } m \text{ access points between reference points and the target location is considered as the most probable estimated location.}
\end{align*}
\]

In addition, machine-learning-based fingerprinting techniques have also been studied for improving the quality of location estimation in complex real environment, including Bayesian modeling, k-nearest-neighbor estimation, support vector machine, neural networks and so on [2, 31, 42, 4].
Figure 2: A geoprocessing-based framework of indoor positioning and tracking.

Figure 3: A geoprocessing workflow for generating Wi-Fi signal surface from samples. Yellow squares represent spatial analysis and data conversion tools while green ellipse shapes represent intermediate data layers and outputs.

standard geospatial Web services. Figure 4 shows a Wi-Fi surface with 1m*1m grids generated by the spatial interpolation process. Note that a grid location in this surface is associated with an array of signal strength from a list of available Wi-Fi access points.

Step 3: The developed mobile app can estimate a user’s indoor location based on a received Wi-Fi signal array $[RSSI(AP)]$. More detailed algorithms will be discussed in the following section. Moreover, we integrate the Firebase\(^4\) real-time database for storing and synchronizing cross-platform app’s indoor positioning data for enabling indoor mobility tracking.

3.2 Positioning Algorithms

Several positioning algorithms have been developed for

\(^4\)https://www.firebase.com/
Wi-Fi positioning [42, 17, 44, 7] and can be categorized as: geometric techniques, statistical methods, fingerprinting and machine learning algorithms. We introduce another algorithm based on image filtering technique and then integrate it with two widely used algorithms: Trilateration and k-nearest-neighbor in signal space (K-NNSS) [2] into our indoor positioning system relying on a Wi-Fi surface which keeps at least three image pixels that meets the multi-value filtering requirements.

As shown in Figure 5, we propose a heuristics-based indoor positioning algorithm in the online mode. It contains three main phases. First, a mobile phone device scans the surrounding environment and gets available Wi-Fi access points with their RSSI value array: \( \text{RSSI}(AP_s) \). Second, a multidimensional-value filtering technique is employed to extract Wi-Fi surface pixels on which the RSSI values on the surface across all available access points are within the signal strength ranges, which can be expressed as follows:

\[
\text{Loc} \in \exists \text{Loc}(j) \left\{ \text{RSSI}(AP_s), -\delta \leq \text{RSSI}(\text{Loc}_j, AP_s) \leq \text{RSSI}(AP_s), +\delta, |\forall i \in (1, 2, ..., N) \right\}
\]

(3)

It assumes that there exist one or several pixel locations \([\text{Loc}_j]\) on the surface such that its corresponding received signal strength is within a constant \(\delta\) deviation away from the current RSSI for each available Wi-Fi access point \(i\) (see Figure 6). Third, depending on the resulting number \(N\) of pixels left after the image filtering step, the program determines which positioning approach should be further triggered. If \(N = 0\), the device needs to re-scan environmental Wi-Fi access points until get updated RSSI values and then repeat previous processing steps; if \(0 < N < 3\), the algorithm will simply return the mean coordinate of the remained surface pixels as the estimated location; if \(N = 3\), the geometric method that utilizes the aforementioned propagation model and trilateration positioning approach (see Section 2.2 for more details about this method) is chosen, and the locations of three Wi-Fi access points with largest RSSIs are taken as the reference points for trilateration; if \(N > 3\), the k-nearest-neighbors algorithm is chosen to approximate the indoor location.

\(K\)-nearest-neighbors is a popular method used for classification and regression [1], in which an object is grouped to the most class by a majority vote of its k-nearest neighbors. In the K-NNSS algorithm, it first calculates the distances between the current observed vector of RSSI values and all RSSI vectors in the signal space.

\[
dist_j(SS_r, SS_o) = \sqrt{\sum_{i=1}^{m} (SS_r(i, j) - SS_o(i))^2}
\]

(4)

where \(SS_r(i, j)\) represents the RSSI of the \(i^{th}\) Wi-Fi access point at a known location \(j\) on the Wi-Fi signal surface, and \(SS_o(i)\) is the measured RSSI of the same Wi-Fi access point \(i\) at the current unknown location.

After calculating all the signal-strength distances with respect to all locations on the signal surface, a set of \(k\) locations with smallest distances is chosen. The chosen value of \(k\) varies in different realistic environments because of changes in available Wi-Fi access points, and a calibration process can help to derive an appropriate value. We choose \(k=5\) in empirical studies. Then the current unknown location \((lat_o, lon_o)\) can be approximated by averaging the coordinate information of each candidate location \((lat_j, lon_j)\) from those chosen k-nearest-neighbor locations in the signal space as follows:

\[
(lat_o, lon_o) = \frac{1}{k} \sum_{j=1}^{k} (lat_j, lon_j)
\]

(5)
or by taking the spatial weights (distance) into consideration, which is the chosen algorithm in developing our indoor-positioning mobile application:

\[
(lat_o, lon_o) = \frac{\sum_{j=1}^{k} \frac{1}{\text{dist}(SS_r, SS_o)} \cdot (lat_j, lon_j)}{\sum_{j=1}^{k} \text{dist}(SS_r, SS_o)}
\]

where \(\text{dist}(SS_r, SS_o)\) is the previously calculated distance between two RSSI vectors. This algorithm is actually consistent with the inverse-distance weighted interpolation in spatial analysis, which estimates the grid-cell attribute value in a raster from a set of sample points by weighting sample-point values so that the farther a sampling point is away from the target cell being evaluated, the less weight it contributes to the calculation of the target cell’s value. The method is inspired by Tobler’s first law of geography: “Everything is related to everything else, but near things are more related than distant things.” [35].

4. IMPLEMENTATION AND RESULT

Based on the geoprocessing framework and workflow introduced above, we first developed a Wi-Fi scanning mobile application to scan and collect the available Wi-Fi access points inside a building of Esri campus (Figure 7a). In the experiments, we selected 18 Wi-Fi access points in a building and collected at least 30 sampling RSSI values for each Wi-Fi access point (Figure 7b). Then the geoprocessing system was built for constructing the signal raster surfaces with 1-meter grid spatial resolution for all available Wi-Fi access points. In the online phase, the indoor position of the user can be estimated via raster multi-value attribute filtering on Wi-Fi signal surfaces and above introduced heuristics-based positioning algorithm. As shown in Figure 7c, when one mobile device was put at the location \(Q\) the red circle represented the estimated position based on our proposed method while the blue circle was the GPS location which still located on the outside of this building. After conducting several rounds of experiments, we got about 2-meter positioning accuracy in average in this empirical study. The user’ locations were also automatically sent to a real-time database (i.e. Firebase) for indoor mobility tracking.

5. DISCUSSIONS

Indoor positioning is a challenging problem and requires a lot of interdisciplinary research. The biggest advantage of Wi-Fi positioning technology is that it can utilize existing wireless-network infrastructures without extra hardware costs as Bluetooth beacons or RFID tags do need. However, it also has some limitations when applying in complex indoor environments:

First, the RSSI values vary dynamically and are sensitive to indoor environments because of signal multi-path fading, path loss, and even the direction to which human bodies are facing. It needs a lot of work in the offline sampling process to ensure that we can get a good approximation of Wi-Fi signal strength database at difference reference locations and surfaces with respect to the spatial structure.

Second, by comparing different positioning algorithms, fingerprinting methods based on machine learning techniques usually get a better position estimation than distance-based trilateration model in real-world complex environments. But fingerprinting methods need database searching and comparison computation processes in the online mode, which takes long time if it involves large-size multidimensional signal surfaces. When choosing the trilateration method, the parameters in the propagation model for converting signal strength to distance values need to be well calibrated. Otherwise sometimes no good matching results can be returned.

Last but not least, by applying the spatial analysis and interpolation methods, it helps to approximate the spatial distribution of signal strength in real-world indoor environments. But it also needs to try different spatial interpolation methods and parameter fitting processes for choosing the best one for generating good signal surface representations.

6. CONCLUSIONS AND FUTURE WORK

A practical WLAN-based positioning system has to be adaptable to the variations of indoor environmental dynamic factors. In this work, we propose a novel Wi-Fi indoor positioning and tracking framework which employs the spatial analysis and image processing techniques. The Wi-Fi surfaces can be dynamically constructed and updated, and thus help to address the challenges of signal spatial heterogeneity and environmental variations. In addition, we don’t need to collect a very large pool of reference points for RSSI sampling comparing with other signal-strength database construction methods in the offline mode, although it is still necessary to get some Wi-Fi signal-strength samples from indoor reference points. Instead, the weighted spatial interpolation method can help to reduce the field sampling burden. A mobile indoor positioning application has been developed as a proof of concept. Based on the experiments we conducted at Esri campus, this method can achieve about 2-meter positioning accuracy. Since it falls into the general category of “fingerprinting”, one can argue that the closer we can approximate to the actual environmental signals, the more accurate we can locate the user. The proposed methodology and theoretical framework can guide engineers to implement cost-effective indoor positioning infrastructure using Wi-Fi in other campuses, and thus offer insights on future smart campus applications. The spatial analysis and geoprocessing workflow may also bring the attention of GIScientists to make more efforts to conquer the indoor positioning challenge.

In future work, we aim to take the signal directionality into consideration and try more advanced spatial interpolation and clustering methods to achieve better accuracy from meter to decimeter. In addition, a crowdsourcing Wi-Fi signal collection and sharing platform will be developed to engage more users’ contribution for advancing indoor positioning technology in our campus. Last but not least, there is a growing trend towards integrating multi-sensors that are available in the mobiles phones, including speaker, accelerometer, gyrometer, barometer etc., to improve indoor positioning accuracy and to semantically label indoor room types. It is also fit into our future investigation into indoor positioning and ambient sensing technologies.

7. REFERENCES

Figure 7: Screenshots of the indoor positioning app using Wi-Fi scanning and spatial analysis techniques. (a) Wi-Fi scanning results at a given location; (b) available Wi-Fi access points in the building (green markers); (c) online-mode indoor positioning (red dot).


A. Motley and J. Keenan. Personal communication.

E. Mok and G. Retscher. Location determination using.

R. Mautz. Overview of current indoor positioning.


Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta. Wi-Fi fingerprinting versus Wi-Fi trilateration.


