A three-dimensional pattern recognition localization system based on a Bayesian graphical model

Abdulraqeb Alhammadi, Fazirulhisyam Hashim, Mohd. Fadlee A Rasid and Saddam Alraih

Abstract
Access points in wireless local area networks are deployed in many indoor environments. Device-free wireless localization systems based on available received signal strength indicators have gained considerable attention recently because they can localize the people using commercial off-the-shelf equipment. Majority of localization algorithms consider two-dimensional models that cause low positioning accuracy. Although three-dimensional localization models are available, they possess high computational and localization errors, given their use of numerous reference points. In this work, we propose a three-dimensional indoor localization system based on a Bayesian graphical model. The proposed model has been tested through experiments based on fingerprinting technique which collects received signal strength indicators from each access point in an offline training phase and then estimates the user location in an online localization phase. Results indicate that the proposed model achieves a high localization accuracy of more than 25% using reference points fewer than that of benchmarked algorithms.

Keywords
Localization, Bayesian inference, fingerprinting

Introduction
Currently, positioning systems are a compelling area of research because they are part of the Internet of things technology. The global positioning system is commonly used in outdoor environments, that is, typically with an unobstructed line of sight (LOS) from a receiver to a satellite. However, it does not function well in indoor environments, given the multipath effect and non-LOS (NLOS) between the transmitter and the receiver. In recent years, indoor positioning issues have received considerable research attention, given the extensive use of wireless local area networks (WLANs) in most indoor environments, such as shopping malls, hospitals and universities. Thus, WLAN-based systems based on time of arrival (TOA) and angle of arrival (AOA) approaches have become extensively used for indoor localization. However, these types of systems require a directional antenna to measure TOA or AOA. The addition of a directional antenna may increase system complexity.

The well-known RADAR is the first deterministic fingerprinting algorithm based on available signal strength measurements. This system mainly depends on the quality of received signal strength indicators
(RSSIs) from access points (APs) which directly affects system accuracy.
By contrast, RSSI-based localization systems do not require angle measurements and additional hardware. Several indoor localization technologies require dedicated infrastructures, such as Wi-Fi, ultrasound signal, ZigBee and ultra-wideband. These technologies have high requirements for the environment and require additional equipment, thereby providing them and other proposed systems with a high level of complexity and inferior accuracy. Moreover, these technologies mainly focus on two-dimensional (2D) planes, whereas three-dimensional (3D) environments are more complicated and significantly increase computational complexity. Research on 3D localization systems has become more realistic than 2D localization systems.

This work proposes an off-the-shelf 3D Bayesian graphical model (3D-BGM) based on radio frequency (RF) fingerprinting technique to predict user location with high accuracy and minimal reference points (RPs). The major contributions of this work can be summarized as follows:

1. Performs long-time analysis of received signal strength (RSS) data to investigate its effect on the localization system;
2. Proposes the 3D-BGM based on the RF fingerprinting technique;
3. Validates the analytical results using the 2D Madigan model.

The remainder of this article is structured as follows. Section ‘Related work’ reviews some related works on indoor localization systems. Section ‘Indoor localization system’ introduces the proposed model and describes the system operation. The experimental design and results are discussed in section ‘Experimental setup and results’. Finally, the conclusion drawn from the current work and brief discussion of future work are presented in section ‘Conclusion’.

Related work

In this section, we present common indoor localization algorithms divided into two approaches, namely, deterministic and probabilistic.

Deterministic technique

The deterministic technique mainly depends on metrics between an online phase and an offline radio map phase. This technique benefits from the K-nearest neighbor (KNN) algorithm when determining the nearest value by comparing online measurements and radio map fingerprints using the Euclidean distance matrix. The estimated location can be ascertained in the convex hull of the K-RPs with the least distance. The KNN algorithm is further improved through the weighted KNN algorithm which assigns weights to each RP in the radio map fingerprint. P Jiang et al. focused on a localization technique based on important APs of Wi-Fi fingerprints; the localization technique relies on APs with the strongest RSS. The authors concluded that the proposed technique reduces the range of fingerprint matching and improves accuracy without pre-knowledge of the building structure. However, the Wi-Fi signals suffer due to the LOS link between APs and mobile devices (MDs) and the effectiveness of the multipath signal propagation. In some cases, localization accuracy depends on radio map fingerprints, where a large radio map leads to reduced localization error and significantly increased computational time. Wang et al. proposed a novel localization scheme based on curve fitting and location search. In their work, an entire area was divided into subareas and fingerprints were created for each subarea. The results demonstrated that the proposed scheme improves localization accuracy by approximately 20% in comparison with other traditional indoor localization algorithms.

In some previous studies, fingerprinting method is implemented on the basis of sparse signal-processing techniques which are considered new-fangled approaches in the deterministic technique. The indicator of the predicted location is considered a vector, where only one or a small subset of indices are non-zero. Therefore, the localization problem can be considered a minimization problem that can be solved by the sparse position vector based on online and offline data. In some recent studies, channel state information (CSI) is used to estimate the location of RSSI, given the high variability of RSS values over time and the multipath effects in indoor environments. However, CSI-based localization requires the use of modified drivers or software-defined radio platforms. This concept makes the CSI-based localization unsuitable for certain environments which require no additional hardware (off-the-shelf) to the end-user.

In another work, the authors proposed an algorithm based on a trilateration method to estimate the location of MDs (smartphones). The propagation paths, such as LOS or NLOS, should be determined to estimate the location of the MDs. The proposed algorithm obtained reduced localization error by up to 1.2 m using three APs. However, it is suitable for small areas (e.g. a small room) where it works with three APs. Thus, it will have a high localization error for large areas. However, Yang et al. improved the conventional trilateration method by considering the greedy algorithm to utilize all effective APs. Conversely, Tuan D Vy and Yoan Shin also improved the triangulation method for
adaptation in large-scale areas, where poor accuracy is probably obtained due to calculation error and path loss propagation.18

Probabilistic technique
The probabilistic technique depends on the statistical likelihood of RSS data at different locations in the environment. The predicted location can be given by acquiring the conditional probability posterior of the prior probability and likelihood function. However, this technique requires an accurate statistical representation of RSS which increases the computational complexity of the system.

Probability distributions are early contributions to the probabilistic approach used for localization algorithms.19 It depends on a set of the strongest APs that have a high probability to cover the entire area. Luo et al.20 applied the Gaussian process–based approach to model the signal distribution of APs and characterize radio map fingerprints in the entire indoor environment. The estimated localization error depended on the fingerprint sampling and error estimation algorithm.

Several sophisticated probabilistic techniques, such as principal component localization,21 conditional random field22 and the Bayesian network, have been studied. Madigan et al.23 introduced a 2D Bayesian system based on the probabilistic approach in which the number of samples is generated from the posterior distribution using Gibbs sampling to predict the users' location based on the maximum posteriori. However, this system utilizes numerous RPs to reduce localization error. Nascimento et al.24 proposed a localization system based on the RF fingerprinting technique. This system used Bayes inference to locate a target in 3D indoor environments. Nevertheless, the majority of the proposed systems are computationally expensive, and the numbers of RPs are overly large. Some systems require additional hardware support associated with MDs. Gu et al.25 introduced a localization scheme with the ability to reduce the measurement effort to construct an offline radio map. The proposed scheme minimized a localization error using only 5% of the collected data. Konstantinos and Richard26 presented a localization scheme based on domain sampling. In the work, the proposed scheme achieved a low localization error of less than 1 ft by a general-purpose solver.

Indoor localization system
This section introduces the Bayesian model–based probabilistic approach and describes the proposed 3D localization model in detail. Furthermore, it discusses the concept of operation of the localization system.

Proposed 3D localization model
The proposed model is based on the Madigan model that only supports 2D environment where it extended work from our previous works.4,27 The proposed model is called 3D-BGM and is an advanced version of the Madigan model, which is designed to support 3D indoor localization systems. Figure 1 illustrates the proposed indoor localization system based on the 3D-BGM. The proposed system consists of five main nodes, namely, AP coordination \((x_i, y_i, z_i)\), user location \((X_j, Y_j, Z_j)\), Euclidean distance \((D_{ij})\), RSS \((S_{ij})\) and testbed dimension \((l, w, h)\).

The proposed model is developed using an OpenBUGS28 tool that uses a visual tool to create graphical models. OpenBUGS refers to Bayesian inference using the Gibbs sampler. The proposed graphical model consists of four stages of representative nodes, which are defined as follows.

First stage. The user location predicted at any point is bounded by the testbed dimension which is considered a uniform distribution. It is defined as

\[
X_i \sim U(0, l) \quad (1)
\]

\[
Y_i \sim U(0, w) \quad (2)
\]

\[
Z_i \sim U(0, h) \quad (3)
\]
where \( l, w \) and \( h \) denote the length, width and height of the testbed dimension, correspondingly, and \((X_i, Y_i, Z_i)\) represent the user’s location at any point of the \(i\)th, which is bound by the testbed dimension \((l, w \text{ and } h)\).

**Second stage.** The distance of an unknown location at any point in the testbed can be expressed by the Euclidean distance \(D_{ij}\)

\[
D_{ij} = \log(a + \sqrt{(X_i-x_j)^2 + (Y_i-y_j)^2 + (Z_i-z_j)^2})
\]

where \(D_{ij}\) represents the distance between AP coordinates \((x_j, y_j\) and \(z_j)\) and users' location \((X_i, Y_i\) and \(Z_i)\). The value of \(a\) is assumed to be 1 to evade the invalid arguments of the logarithm function.

**Third stage.** RSS is defined as a normal distribution that has mean and variance equal to the regression model of independent variables \(b_{\theta_0}\) and \(b_{\theta_1}\) and \((\tau_{\theta_0}\) and \(\tau_{\theta_1}\), respectively

\[
\begin{align*}
S_i & \sim N(b_{\theta_0} + b_{\theta_1}\log D_{ij}, \tau_i) \\
\theta_0 & \sim N(b_0, \tau_{\theta_0}), b_0 \sim N(0.001, \tau_{\theta_0} - \gamma(0.001, 0.001)), \\
\theta_1 & \sim N(b_1, \tau_{\theta_1}), b_1 \sim N(0.001, \tau_{\theta_1} - \gamma(0.001, 0.001))
\end{align*}
\]

The RSS is measured at the \(i\)th user location and \(j\)th AP location. \(S_{ij}\) is the normal distribution defined as \(S_{ij} \sim N(\mu, \tau)\). The regression model is assigned as the mean of the normal distribution of \(S_{ij}\). It consists of four parameters (i.e. \(b_0, b_1, b_2\) and \(b_3\)) and one independent variable \(D_{ij}\).

**Fourth stage.** The initial parameters are the normal distributions \(b_{\gamma} \sim (\mu, \tau)\) that carry any arbitrary values used to start the burning-in generating samples only in the initial stage. The parameters are defined as follows

\[
\begin{align*}
b_{\theta_0} & \sim (\mu_0, 0), b_{\theta_1} \sim (\mu_1, 1), b_{\gamma} \sim (\mu_{\gamma}, \tau_{\gamma}) \\
v & \sim (b, \tau) \text{ and } v \sim (0.001) \\
\tau_{\gamma} & \sim (0.001, 0.001)
\end{align*}
\]

The Bayesian probability interprets the theorem expression which infences user location (posterior) based on a radio map (prior). The posterior of conditional probability is equal to the product of prior probability and likelihood function

\[
\text{Posterior} = \text{prior} \times \text{likelihood}
\]

The Gibbs sampling algorithm is used to draw samples from the highly complicated probability based on prior distribution.\(^{29}\) It draws samples \(s_k^{(i+1)}\) from the conditional probability, given the initial value \(s^{(0)}\).

\[
\begin{align*}
\text{Gibbs sampling algorithm:} \\
& \text{Initialize the initial values } s^{(0)} = (s_1^{(0)}, \ldots, s_K^{(0)}) \\
& \text{for loop } (i: 1; N) \\
& \quad \text{Drawing samples of } s_{k}^{(i+1)} \text{ from } P(s_{k}^{(i+1)} | s_1^{(i)}, \ldots, s_{k-1}^{(i)}, s_{k+1}^{(i)}, \ldots, s_K^{(i)}) \\
& \quad \text{... Drawing samples of } s_{k}^{(i+1)} \text{ from } P(s_{k}^{(i+1)} | s_1^{(i+1)}, \ldots, s_{k-1}^{(i+1)}, s_{k+1}^{(i+1)}, \ldots, s_K^{(i+1)}) \\
& \quad \text{Return the values } (s_1^{(i)}, s_2^{(i)}, \ldots, s_K^{(i)})
\end{align*}
\]

Consequently, numerous samples which represent the posterior distribution of the unknown location will be drawn. The first Bayesian system mainly consists of nodes which represent the variables related to some parameters that can be used for the localization system. These parameters include the distance between MDs and APs.

**Fingerprinting localization operations**

The proposed 3D-BGM is a device-free localization system that does not require any additional equipment to be used in addition to the available APs and MDs. Figure 2 displays the process of the RF fingerprinting technique which consists of two phases described as follows.

**Offline phase.** This phase is also called the data collection phase and is responsible for collecting samples of RSSI fingerprints (known as RPs) using an MD that supports Wi-Fi technology. The user stands with a device at the location of interest within the testbed and collects RSSI samples from all available APs at time \(t_m\), where \(m = 1, 2, \ldots, M\). The collected RPs associated with RSSI can be expressed as follows

\[
i = (x_i, y_i, z_i) \text{ and } i = 1, 2, \ldots, N \]

\[
\chi = (\chi_1, \ldots, \chi_N) \begin{bmatrix} \chi_1^1 & \cdots & \chi_N^1 \\ \vdots & \ddots & \vdots \\ \chi_1^K & \cdots & \chi_N^K \end{bmatrix}
\]

where \(i\) represents the set of RPs in the Cartesian coordinates at any point in the experimental testbed, \(N\) is the number of collected RPs stored in the radio map at different locations in the area of interest, \(\chi\) denotes the collected samples of RSS at each RP with \(\chi_i = [\chi_i^1, \ldots, \chi_i^K]^T\) and \(\chi_i^k = \sum_{m=1}^{M} t_m / M\) and \(K\) represents the number of available APs in the testbed area. The average values of RSSI from each AP at different locations are used to construct the radio map.
**Online phase.** This phase is responsible for receiving samples of RSS from available APs and comparing the current RSS samples with collected data in the radio map constructed during the offline phase to estimate the unknown location. The MD receives $Y$ online RSS observations, which contain the current RSS from each AP $Y = \frac{1}{2^j} u_1, \ldots, \frac{1}{C_1} u_j$ at any unknown location. Subsequently, the current RSS is compared with the radio map using the fingerprinting technique. Finally, the mobile location is estimated by inferring its coordinates among the optimal matches on the radio map $c_{Ox}, c_{Oy}, c_{Oz} = \text{argmin}_i D_{ij}(\theta)$. The system accuracy measures the overall performance of the proposed algorithm or models for location prediction which depends on the calculated localization error $E_i$ of each training point in the system. The localization error can be defined by the Euclidean distance $E_i$

$$E_i = \sqrt{(O_x - \widehat{O_x})^2 + (O_y - \widehat{O_y})^2 + (O_z - \widehat{O_z})^2}$$

where $(O_x, O_y, O_z)$ and $(\widehat{O_x}, \widehat{O_y}, \widehat{O_z})$ correspond to the actual and estimated locations for the MD of the $i$th RPs, respectively. The overall system accuracy refers to the average of the overall localization error which can be expressed as

$$\text{System accuracy} = \frac{\sum_{i=1}^{q} E_i}{q} \times 100\%$$

where $q$ denotes the number of training points used to test the proposed model. Algorithm 1 demonstrates the steps for estimating the user location based on the proposed 3D-BGM. $\xi$ denotes the number of burning-in samples which will be discussed in the next section.

**Experimental setup and results**

**Experimental design**

To test the performance of the proposed 3D-BGM, we conducted an experiment in an indoor environment with a dimension of $50 \times 22$ m$^2$. Four APs were used to collect the RSS fingerprints along the corridor which contained 50 RPs from each AP, as exhibited in Figure 3. The black dots represent the RPs along the corridor. Tables 1 and 2 summarize the specifications of the testbed, APs and MDs. Wi-Fi scanner software was used to scan the available APs and collect data, such as RP coordinates, media access control (MAC) address, service set identifier (SSID), channel, RSS and timestamp for each selected AP. In this work, 30 samples (1-s intervals) were collected in a 360 degree rotation for each RP and each AP along the corridor. Two experiments were performed with time durations at the same place using the same APs and MDs. The gap between these experiments was 3 years to study the effect of RSS properties and their impact on system accuracy over a long period. The data collection process for different time durations can be found in a previous paper.\(^{30}\)

**Algorithm 1**

1. Input: initializing the input parameters: equations (1)–(5)
2. Output: estimating user locations and obtaining system accuracy
3. while $t < q$, do
4. If $I = \xi$
5. while no. of iteration $\leq k$ do
6. Apply Gibbs sampling to draw samples
7. $s^{k+1}_i$ from $P(s_i|s_j^{(i+1)}, \ldots, s_k^{(i+1)})$
8. $\{s^{(1)}, s^{(2)}, \ldots, s^{(k)}\}$
9. end while
10. else if
11. Update $b_0, b_j, \tau_0$ and $\tau_1$
12. end if
13. Calculate the average of generated samples for the estimated location for $(X, Y$ and $Z$) within testbed dimension.
14. Calculate the localization error using equation (8)
15. end while
16. Calculate the overall system accuracy using equation (9)
Performance evaluation

Impact of the RSS across times. To study this effect, an experiment was conducted to collect the RSS data along the corridor during different timeframes with a gap of 3 years. The datasets were conducted at the same place with the same number of APs using the same MD. Figure 4(a) and (b) demonstrates the average value of RSS for 50 RPs in the first and second datasets, respectively. The high and low values of RSS were obtained when the location of the MD was close to and far from the APs, respectively, for both datasets. All RSS reading from both datasets obtained the same average value in the middle of the corridor (intersection area at 26 m). This result was because the MD was the midpoint between all APs in the testbed. In addition, RPs possess a unique set of RSS which makes the predictions of unknown locations using fingerprint-based localization techniques the best choices. In our previous work, an investigative study was conducted by introducing three types of RSS data that might influence the location prediction for each training point. The study concluded that changes in environment structure must be considered to predict an unknown location with high accuracy.

Figure 5 illustrates the impact of RSS reading during a long gap period. The deviation in the RSS reading was due to the multipath effect and attenuation caused by changes in the testbed structure and the movements of people. These factors led to fluctuating signal strengths at different time durations. AP2 recorded the highest value standard deviation among the APs.

RSS distribution. RSS distribution does not constantly have Gaussian or asymmetric properties due to changes in signal levels over time. However, the distribution of RSS is defined as a non-Gaussian distribution, considering the frequently different RSS means and modes. The distribution of RSS is further defined as a normal or Gaussian distribution when a similarity exists in RSS mean and median readings. Thus, the distribution of RSS is difficult to model and fit to a particular distribution due to the complexity of the radio propagation of

Table 1. Testbed specification.

| Parameter                  | Specification |
|----------------------------|---------------|
| Testbed dimension (m²)     | 50 × 22       |
| Height (m)                 | 2.65          |
| Number of APs              | 4             |
| Structural Wall types      | Concrete, glass and plasterboard |
| Internal wall thickness (cm)| 15            |
| External wall thickness (cm)| 20            |

AP: access point.

Table 2. Specifications of used MD and APs.

| Device          | Parameter     | Value         |
|-----------------|---------------|---------------|
| MD (laptop)     | Brand         | Acer          |
|                 | WLAN card     | Atheros AR5007EG |
|                 | Processor speed (GHz) | 2.5          |
|                 | RAM (GB)      | 4             |
| APs             | Model         | Linksys-Cisco |
|                 | Operating frequency (GHz) | 2.4          |
|                 | Transmit power (dBm) | 18            |

MD: mobile device; APs: access points; WLAN: wireless local area network.
indoor environments. To investigate the RSS distribution in the testbed, RSS measurements were collected from AP1 for 10 min (one sample per second), with a distance of 8 m between the AP and the MD. The collected data were divided into three time intervals (i.e. 2, 5 and 10 min) to study the effect of each part separately. Figure 6 plots the histogram of the RSS distribution for the three time intervals. We observed that RSS behaves similar to a normal distribution in the second and third time intervals due to the similarity of the RSS mean and median. However, the RSS was a non-Gaussian distribution when the first time interval was compared with the second or third time intervals considering the different means and medians. Overall, the RSS distribution in this particular test was defined as non-Gaussian and asymmetric.

**RSS stationary.** The stationary test used to investigate the mean and variance of RSS does not change over different time intervals. To conduct this type of RSS property test, the collected data were divided into two parts, where each part contained 300 samples (one sample per second) of RSS. The stationary decision of RSS was based on two conditions that RSS must satisfy. The first condition is that the mean and variance of the RSS must remain the same over time. The second condition is that its autocorrelation function should have the same shape during the time interval. Figure 7 demonstrates the method for determining that the RSS has consistent mean and inconsistent variance values for each part of the time interval. Therefore, the RSS process was considered non-stationary in this case due to the failure of the first condition of the stationary test. Figure 8(a) and (b) displays the same shapes of the autocorrelation function for each part. The similarity in shapes indicates that the second condition of the stationary test was satisfied. In summary, the RSS random

![Figure 4](image4.png)

**Figure 4.** Average RSS with respect to APs for different time durations. (a) First and (b) second RSS readings.

![Figure 5](image5.png)

**Figure 5.** Effect of RSS reading during a long gap period.

![Figure 6](image6.png)

**Figure 6.** Histogram of the RSS distribution at different time intervals: (a) 2 min, (b) 5 min and (c) 10 min.
process is non-stationary due to failure of the first condition, although it satisfied the second condition.

User’s body effect. The RSS is affected by the presence of user body due to the multipath phenomena in indoor environments. The effect of user’s body must be investigated and considered before designing an indoor localization system. Typically, the user carries the MD for collecting data in a particular area. An experiment was conducted to study the significant effect of this parameter on system accuracy. Samples of RSS were collected from AP1 for 5 min with an 8-m distance from AP1 in two phases (user presence and no user presence). Figure 9 exhibits the effect of the user’s body on RSS. The result showed that the existence of the user slightly reduces the RSS mean by $-2.77$ dBm, whereas its standard deviation evidently increased from 4.45 to 6.03. The human body is considered an additional source of inaccuracy that can cause unpredictable fluctuations in RSS. That is, the human body is an effective absorber of 2.4 GHz of WLAN radio signal because it is composed of 70% water that causes degradation in RF performance.34

Inferencing user location

The proposed 3D-BGM was evaluated using OpenBUGS to estimate any point of user location. Table 3 provides the values of the parameters for location inference used for this work. The burn-in samples refer to the practice of discarding the initially generated samples to eliminate their effect on the posterior inference. To illustrate the generation of samples, the fourth RP ($X = 24$, $Y = 9$ and $Z = 1$) in the radio map was taken as an example to show the inference process. Figure 10 presents a trace of random variables for $X[4]$, $Y[4]$ and $Z[4]$. The two random variables reach a convergence level which signifies that increased numbers of iterations have no significant effects.

The samples were generated using the Gibbs sampler for variables $X[4]$ and $Y[4]$, as depicted in Figure 11. These samples were obtained by running a Markov chain for 100,000 iterations. The execution time required to obtain these iterations was 62 s. Figure 12 illustrates the autocorrelation of the generated samples for random variables $X[4]$ and $Y[4]$. 

![Figure 7. Collected RSS over time (10 min).](image7)

![Figure 8. Autocorrelation function of two time intervals. (a) First and (b) second parts.](image8)

![Figure 9. Effects of user presence and no user presence on the indoor environment.](image9)
Impact of the number of iterations. The number of iterations is an important factor in the positioning accuracy of our proposed model. An increase in the number of iterations leads to an increased probability of the system to estimate the correct user location. However, numerous iterations increase the computational time abruptly. Table 4 displays seven sets of the number of iterations investigated to study their effect on the accuracy of the proposed model. Figure 13 displays the average of the distance error which gradually decreased during the first five sets of iterations (20,000–100,000 iterations). However, no significant improvement was noted for the last three sets of iterations (100,000–140,000 iterations). Conversely, the model reached a convergence level at 100,000 iterations with an average distance error of 2.9 m. Thus, the optimum choice of the number of iterations for the proposed model is 100,000 because no improvement was observed when the number of iterations increased.

Figure 14 demonstrates the effect of the number of training points used to test the proposed 3D-BGM and the Madigan model using the first and second RSS datasets. Four sets of training points (set1 = 6, set2 = 9, set3 = 12 and set4 = 15) were investigated using both RSS datasets for each model. The localization system achieves high accuracy when the training points are increased for both models. Furthermore, the proposed 3D-BGM and the Madigan model that uses the second RSS dataset reduced the average distance error

**Table 3.** Parameters for inference setting.

| Parameter          | Value       |
|--------------------|-------------|
| Number of chains   | 1           |
| Burn-in samples    | 10,000      |
| Number of iterations | 100,000   |
| Refresh            | 100         |
| Thin               | 1           |
| Inference nodes    | X, Y, Z     |

**Figure 11.** Generated samples using the Gibbs sampler for X[4] and Y[4].

**Figure 12.** Autocorrelation function of the generated samples at the (a) X and (b) Y coordinates.
in comparison with that of the first dataset. However, the proposed model outperformed the Madigan model for all datasets.

Figure 15 exhibits the rate of localization error for different types of datasets using the 2D Madigan model and the proposed 3D-BGM. Clearly, the 3D-BGM outperformed the Madigan model for the first and second datasets. The overall average localization accuracies for the proposed 3D-BGM and the Madigan model were 2.9 and 3.8 m, correspondingly. Moreover, the 3D-BGM achieved high accuracy using only four APs with a small number of RPs in comparison with the Madigan model.

The proposed 3D-BGM was compared with different localization algorithms, such as Hyeon et al.’s model and KNN using new data, in addition to the Madigan model. The Hyeon model is based on a trilateration method that used three APs to estimate the user location in a small area. In this comparison, the same testbed dimension, number of training points and number of APs were used (except for the Hyeon model, where only three NLOS APs were used) for all
algorithms. Figure 16 depicts the comparison between different localization algorithms using various sets of training points (i.e. set1 = 6, set2 = 9, set3 = 12 and set4 = 15). The same specifications of the testbed were used for all testing sets and compared models. The results demonstrate that the proposed 3D-BGM achieves a significant reduction of localization error in comparison with other algorithms for all different sets. The KNN algorithm obtained the lowest average localization accuracy for all tested sets because larger testing points fail to match highly similar locations. The Madigan and Hyeon models recorded better localization accuracies at (using 15 training points) 3.8 and 7.6 m, respectively, than the KNN algorithm. The proposed 3D-BGM achieves an average system accuracy of 2.9 m higher than that with the Madigan, KNN and Hyeon algorithms. In particular, the proposed model improved system accuracy by 25%, 73% and 62% in comparison with the Madigan, KNN and Hyeon algorithms, correspondingly.

Conclusion
This work presented the design, analysis and evaluation of 3D-BGM for indoor localization systems. The 3D-BGM based on the RF fingerprinting technique used available APs already deployed in the environment to estimate user location without additional external devices. The proposed 3D-BGM achieved an overall localization error of 2.9 m using only four APs with a few RPs. This model provided accuracies that are higher by 25%, 62% and 73% than those of the Madigan, KNN and Hyeon algorithms, respectively. In the future, the 3D-BGM will be further enhanced by considering a multi-story building rather than a single-floor unit. This condition will be implemented by adding a new parameter to the proposed 3D-BGM called the ‘floor attenuation factor’, where the RSS attenuates few dBs because it penetrates each floor.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Abdulraqeb Alhammadi https://orcid.org/0000-0003-0778-1869

References
1. Güvenc I and Chong CC. A survey on TOA based wireless localization and NLOS mitigation techniques. *IEEE Commun Surv Tut* 2009; 11: 107–124.
2. Tomic S, Beko M and Rui D. Distributed RSS-AoA based localization with unknown transmit powers. *IEEE Wirel Commun Let* 2016; 5: 392–395.
3. Bahl P and Venkata PN. RADAR: an in-building RF-based user location and tracking system. In: *INFOCOM 2000. Nineteenth annual joint conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 2, Tel Aviv, 26–30 March 2000. New York: IEEE.
4. Alhammadi A, Hashim F, Mohd F, et al. A new indoor localization system based on Bayesian graphical model. In: *International conference on wireless communications, signal processing and networking (WiSPNET)*, Chennai, India, 22–24 March 2017, pp.1960–1964. New York: IEEE.
5. Zafari A, Gkelias F and Leung KA. Survey of indoor localization systems and technologies. Computing Research Repository, 2017, arXiv:1709.01015v1.
6. Han S, Zhao C, Meng W, et al. Cosine similarity based fingerprinting algorithm in WLAN indoor positioning system. In: *2015 IEEE international conference on communications (ICC)*, London, 8–12 June 2015, pp.2710–2714. New York: IEEE.
7. Jiang P, Zhang Y, Fu W, et al. Toward robust indoor localization based on Wi-Fi fingerprint’s important access point. *Int J Distrib Sens N* 2015; 2015: 429104.
8. Wang J, Gao Q, Yu Y, et al. Toward robust indoor localization based on Bayesian filter using chirp-spread-spectrum ranging. *IEEE T Ind Electron* 2012; 59(3): 1622–1629.
9. Alraih S, Alhammadi A, Shayea I, et al. Improving accuracy in indoor localization system using fingerprinting technique. In: *2017 international conference on information and communication technology convergence (ICTC)*, Jeju, South Korea, 18–20 October 2017, pp.274–277. New York: IEEE.
10. Alhammadi A, Alias MY, Tan SW, et al. An enhanced localization system for indoor environment using clustering technique. *Int J Comput Vis Robot* 2017; 7(1–2): 83–98.
11. Wang B, Zhou S, Liu W, et al. Indoor localization based on curve fitting and location search using received signal strength. *IEEE T Ind Electron* 2015; 62(1): 572–582.

12. Feng C, Au W, Valae S, et al. Received-signal-strength based indoor positioning using compressive sensing. *IEEE T Mobile Comput* 2012; 11(12): 1983–1993.

13. Wang X, Gao L, Mao S, et al. CSI-based fingerprinting for indoor localization: a deep learning approach. *IEEE T Veh Technol* 2017; 66(1): 763–776.

14. Zhang Y, Li D and Wang Y. An indoor passive positioning method using CSI fingerprint based on AdaBoost. *IEEE Sens J* 2019; 19(14): 5792–5800.

15. Jo HJ and Kim S. Indoor smartphone localization based on LOS and NLOS identification. *Sensors* 2018; 18: 3987.

16. Yang J, Li Y and Cheng W. An improved geometric algorithm for indoor localization. *Int J Distrib Sens N* 2018; 14: 1–13.

17. Vy TD and Shin Y. iBeacon indoor localization using trusted-ranges model. *Int J Distrib Sens N* 2019; 15: 1–13.

18. Tong H, Deng Q, Zhang T, et al. A low-cost indoor localization system based on received signal strength indicator by modifying trilateration for harsh environments. *Int J Distrib Sens N* 2018; 14: 1–11.

19. Luo C, Cheng L, Chan MC, et al. Pallas: self-bootstrapping fine-grained passive indoor localization using WiFi monitors. *IEEE T Mobile Comput* 2017; 16(2): 466–481.

20. Luo C, Hong H, Cheng L, et al. Accuracy-aware wireless indoor localization. *J Netw Comput Appl* 2016; 62: 128–136.

21. Fang SH and Lin T. Principal component localization in indoor WLAN environments. *IEEE T Mobile Comput* 2012; 11(1): 100–110.

22. Xiao Z, Wen H, Markham A, et al. Lightweight map matching for indoor localization using conditional random fields. In: *IPSN-14 proceedings of the 13th international symposium on information processing in sensor networks*, Berlin, 15–17 May 2014, pp.131–142. New York: IEEE.

23. Madigan D, Einahrawy E, Martin RP, et al. Bayesian indoor positioning systems. In: *Proceedings IEEE 24th annual joint conference of the IEEE Computer and Communications Societies*, vol. 2, Miami, FL, 13–17 March 2005, pp.1217–1227. New York: IEEE.

24. Nascimento HJ, Rodrigues EB, Cavalcanti FR, et al. An algorithm based on Bayes inference and K-nearest neighbor for 3D WLAN indoor positioning. In: *Proceedings of Simpósio Brasileiro De Telecomunicacoes (SBRT)*, Santos, Brazil, 30 August–2 September 2016, pp.398–402.

25. Gu Z, Chen Z, Zhang Y, et al. Reducing fingerprint collection for indoor localization. *Comput Commun* 2016; 83: 56–63.

26. Konstantinos K and Richard MP. Reducing the computational cost of Bayesian indoor positioning systems. In: *2006 3rd annual IEEE Communications Society on sensor and ad hoc communications and networks*, vol. 2, Reston, VA, 28 September 2006. New York: IEEE.

27. Alhammadi A, Alraith S, Hashim F, et al. Robust 3D indoor positioning system based on radio map using Bayesian network. In: *2019 IEEE 5th world forum on Internet of Things (WF-IoT)*, Limerick, 15–18 April 2019, pp.107–110. New York: IEEE.

28. Thomas. OpenBUGS, 2004, http://www.openbugs.net/w/ FrontPage

29. Gelfand AE and Smith AFM. Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 1990; 85(410): 398–409.

30. Alhammadi A, Hashim F, Mohd F, et al. Effects of different types of RSS data on the system accuracy of indoor localization system. In: *2016 IEEE Region 10 symposium (TENSYMP)*, Bali, Indonesia, 9–11 May 2016. New York: IEEE.

31. Kaemarungsi K and Krishnamurthy P. Properties of indoor received signal strength for WLAN location fingerprinting. In: *First annual international conference on mobile and ubiquitous systems: networking and services (MobiQuitous’04)*, Boston, MA, 26 August 2004, pp.14–23. New York: IEEE.

32. Ladd AM, Bekris KE, Rudys A, et al. Robotics-based location sensing using wireless Ethernet. *Wirel Netw* 2005; 11: 189–204.

33. Small J, Smailagic A and Siewiorek DP. Determining user location for context aware computing through the use of a wireless LAN infrastructure. *Project Aura report, Institute for Complex Engineered Systems*, Pittsburgh PA, 2000.

34. Della RF, Pelosi M and Nurmi J. Human-induced effects on RSS ranging measurements for cooperative positioning. *Int J Navig Obs* 2012; 2012: 13.