

## Chapter 4

# Survey-Based Location Systems

The two main primitives in location systems—range and angle-of-arrival—were introduced in [Chap. 2](#). By measuring either of the two between a mobile device and the base stations in a backbone network, the primitives can be respectively triangulated or angulated to an estimated location for the mobile. For indoor systems that require high accuracy and fidelity, the range primitive—when measured through time-of-arrival (TOA) techniques—is the more robust of the two to signal fading, as explained in [Chap. 2](#). The success of time-of-arrival techniques hinges upon the predictable mapping between the TOA and the distance travelled by the signal. The range primitive can also be measured through received-signal-strength techniques, however, in harsh propagation environments, the mapping is very sensitive to fading, which is nondeterministic in nature. Hence, range-based mapping cannot be reliably exploited for RSS systems in such environments.

Yet, received signal strength has been shown to deliver decent accuracy even in indoor environments—when used in survey-based location systems. In survey-based systems, RSS is not mapped to range, but directly to location. The technique is just to assume that because signal loss occurs in the environment—not only due to path loss, but also due to penetration loss and specular effects such as reflection and diffraction—such a mapping exists. Because the mapping is so complex, there is no attempt to explicitly model the received signal strength as a function of location. Instead, the mapping is constructed by observing the “fingerprints” that the RSS “leaves” throughout the environment. From the observed fingerprints, the RSS-location mapping can be reconstructed. In practice, fingerprinting systems associate values of a physically measureable feature to discrete locations throughout a survey area. RSS is the most common feature but, as we shall see, others have become popular recently. Then, a mobile device can estimate its location based on the value it measures during a query. The feature value at a particular location is known as a fingerprint, or signature, because it can be used to identify the location.

One of the earlier and most simple fingerprinting systems, known as RADAR (Bahl and Padmanadhan 2000), is based on the received signal strength. The simplicity of this indoor location system stems from the fact that RSS measurements are readily available in the IEEE 802.11 standards implementation. For outdoor systems, on the other hand, the RSS is measured from cellular towers or satellites. More on cellular systems is discussed in the following chapter. Since base stations (access points) typically have overlapping coverage, the actual feature is the vector of RSS values received from all available base stations. Before the system can be operational, a radio map of the environment must be constructed in a so-called fingerprinting stage. In this stage, a discrete set of  $n_M$  candidate sites  $\mathbf{x}_i, i = 1 \dots n_M$ , for the mobile is selected throughout the survey area a priori. At each site, the received powers from the base stations are recorded and stored in a database. Let  $n_B$  denote the maximum number of base stations from which a mobile device can receive a distinct signal. Then, the signature at location  $\mathbf{x}_i$  is the vector of received powers denoted as  $\mathbf{P}_i = [P_{i1}, P_{i2}, \dots, P_{i,n_B}]^T$ . We refer to an ordered pair composed from a location and its associated signature  $(\mathbf{x}_i, \mathbf{P}_i)$  as a *training pair*.

During system operation—which is known as the localization stage—a vector  $\hat{\mathbf{P}} = [\hat{P}_1, \hat{P}_2, \dots, \hat{P}_{n_B}]^T$  of received powers is measured at the mobile device. RADAR uses the nearest neighbor method as a mapping algorithm from the measured power to the estimated location for the mobile. Specifically, the mobile's location is determined as the location  $\mathbf{x}_c$  of the registered site whose fingerprinted power  $\mathbf{P}_c$  is closest to the measured power  $\hat{\mathbf{P}}$  in terms of some similarity metric in the RSS vector space.

Location fingerprinting systems can be differentiated for the most part by the following two characteristics: (1) the feature selected to fingerprint the sites; and (2) the mapping algorithm to determine the mobile's location. In this chapter, we introduce several fingerprinting techniques. Given its prevalence, we concentrate on the RSS feature in the first part of the chapter. The same techniques, however, apply to other features as well. In the first section, an analytical model of a generic fingerprinting system is presented. The model describes how the salient parameters common to most systems affect their performance. The subsequent section showcases a number of methods to compute the similarity metric for memoryless systems—that is—systems which estimate location based on readings taken at a single time instant. Section 4.3 introduces systems with memory and shows how maintaining some historic path data can enhance location precision significantly. In the remainder of the chapter, we introduce some non-RSS features. Section 4.4 investigates the use of the channel impulse response as an alternative radio frequency signature. Conversely, Sect. 4.5 reports on non-RF features altogether—features which are available from devices such as smartphones, namely sound, motion, and color.

## 4.1 Analytical Models

Besides the selection of the feature (or features) in any fingerprinting system, there are a number of system factors which affect performance, most notably the number of base stations, the number of fingerprinted sites, and the spacing between the sites. Naturally, the harshness of the propagation environment also affects performance. In this section, we describe two analytical models proposed in Kaemarungsi and Krishnamurthy (2004) to investigate these factors. Again, although the proposed models are specific to RSS-feature systems, the principles apply to all types of fingerprinting systems. As in the RADAR system, it is assumed that the vector of received signal strengths from the base stations is used to fingerprint the sites.

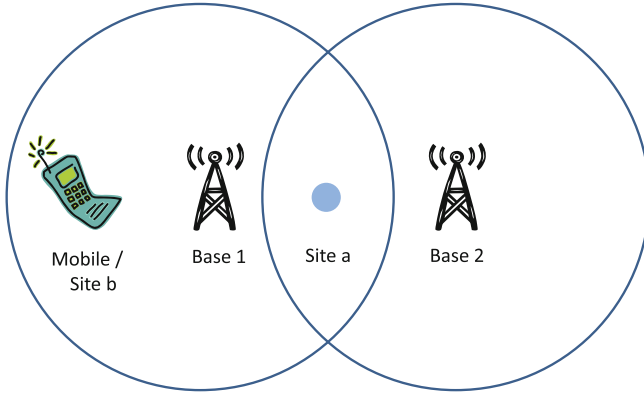
### 4.1.1 A Stochastic Model for the Similarity Metric

A popular similarity metric between the measured power,  $\hat{\mathbf{P}}$ , and fingerprinted power at a particular site indexed as  $i$ ,  $\mathbf{P}_i$ , is the square Euclidean norm in the  $n_B$ -dimensional space (Liu et al. 2007):

$$\begin{aligned} \rho_i &= \|\hat{\mathbf{P}} - \mathbf{P}_i\| \\ &= \sum_{j=1}^{n_B} (\hat{P}_j - \hat{P}_{ij})^2. \end{aligned} \quad (4.1)$$

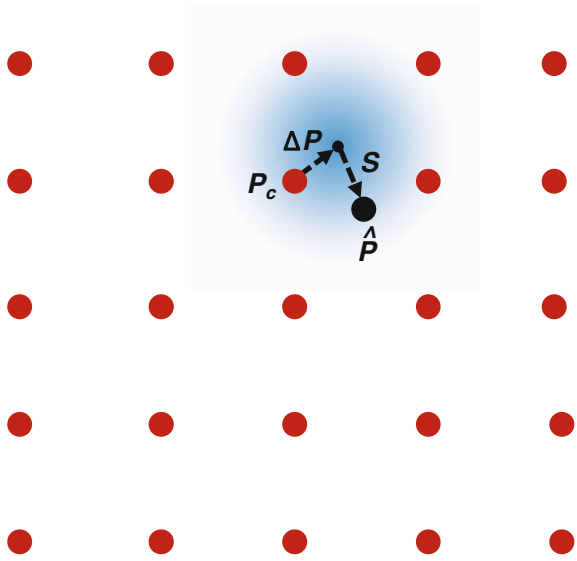
The units for the power are in dBm. In practice, a mobile device may not receive a signal from all  $n_B$  base stations. This is because when a mobile device is far away from a site—especially in large deployment areas—the set of base stations from which it receives may differ from the set registered at the site. In this case, the similarity metric can compare only the signal strengths from the common base stations. So, a penalty term  $\rho_0$  is added to (4.1) instead for each base station which is not common to both sets, where  $\rho_0$  is a system-specific tuned constant. Figure 4.1 illustrates a simple case for which the penalty term is functional. Site a is registered to both stations whereas Site b, since it lies beyond the coverage area of Base 2, is only registered to Base 1. Since the mobile is at Site b, it also cannot receive from Base 2. As such, the similarity metric is computed only from Base 1. Since both sites are equidistant from Base 1, without the penalty term they would have equal similarities; on the other hand, by penalizing Site a because there is no reception from Base 2, the mobile's location can be successfully resolved to Site b.

In the fingerprinting stage, the sites are selected on a square grid throughout the deployment area. The fingerprinted power at a site is the expected value of the received power at the location, i.e., neglecting the stochastic effects of shadowing. The measured power at the mobile device during a location query can be modeled as the sum of three terms. The sum is expressed as



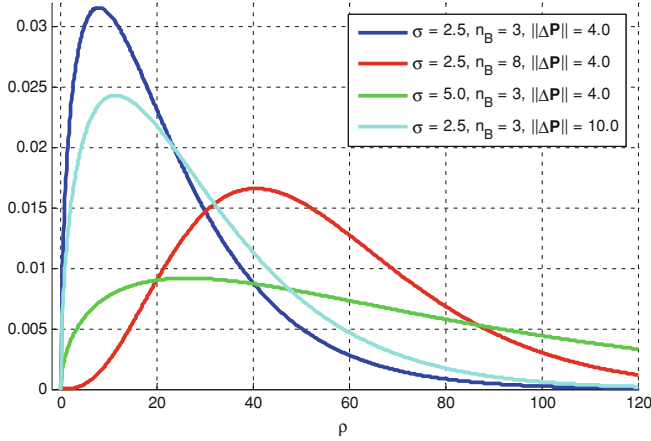
**Fig. 4.1** Site a is within the coverage areas of both base stations while Site b is only within the coverage area of Site b. Although both sites are equidistant from Base 1, because the mobile only receives from Base 1, the mobile’s location can be resolved to Site b

**Fig. 4.2** A  $5 \times 5$  grid of fingerprinted sites (red). The actual location of the mobile device is shown in black. The expected location of the mobile is in the middle of the radial pattern, however, due to shading, the mobile may be found anywhere. The probability of finding the mobile decreases as the radial pattern fades away



$$\hat{P} = P_c + \Delta P + S. \tag{4.2}$$

The first term is the signature power of the site which is most similar to the measured power. This site is indexed as site  $c$  because it indicates the correct estimate for the mobile’s location. The second term is the offset between the mobile’s expected power and the fingerprinted power at site  $c$ . And the third term is the fluctuation of the signal due to shadowing. Figure 4.2 illustrates these three components. Shown is a  $5 \times 5$  grid of fingerprinted sites (red). The expected location of the mobile is at the center of the radial pattern. Due to shadow fading,



**Fig. 4.3** The non-central Chi-squared distribution represents the pdf of the similarity metric  $\rho_c$ :  $\sigma$  is the standard deviation of the shadow fading parameter,  $n_B$  is the number of base stations in the system, and the parameter  $\|\Delta\mathbf{P}\|$  is proportional to the grid spacing between the fingerprinted sites in the area

however, the mobile (black) may be found anywhere. The probability of finding the mobile decreases as the radial pattern fades away.

Since the power is measured on a logarithmic scale, the  $n_B$  individual components of  $\mathbf{S}$  are distributed normally as  $S_j \sim \mathcal{N}(0, \sigma)$  [see Chap. 3]. By substituting Eq. (4.2) into (4.1), the latter then reduces to  $\rho_c = \|\Delta\mathbf{P}\| + \|\mathbf{S}\|$ . The resultant distribution for the similarity metric of the correct location is the non-central Chi-square probability density function (pdf) with  $n_B$  degrees of freedom:

$$f_{\rho_c}(\rho) = e^{-\frac{(\|\Delta\mathbf{P}\| + \rho)}{2\sigma^2}} \frac{1}{2\sigma^2} \left(\frac{\rho}{\|\Delta\mathbf{P}\|}\right)^{\frac{(n_B-2)}{4}} J_{\frac{n_B-2}{2}}\left(\frac{\sqrt{\|\Delta\mathbf{P}\|\rho}}{\sigma^2}\right), \quad \rho \geq 0, \quad (4.3)$$

where  $J_\vartheta(\cdot)$  is the  $\vartheta$ -th order Bessel function of the first kind.

The effects of  $\sigma$ ,  $n_B$ , and  $\Delta\mathbf{P}$  on the similarity metric are illustrated in Fig. 4.3. Naturally, the pdf spreads out as the amount of shadowing, represented by  $\sigma$  increases. Adding base stations to the system magnifies this effect since there will be more shadowing components in  $\mathbf{S}$ . The latter phenomenon is captured in Eq. (4.3) through the associated parameter  $n_B$ , which spreads the curve out yet further. Although with additional stations the similarity metric is more susceptible to shadowing, the enhanced identifiability that the stations bring to the sites delivers better performance overall. This is highlighted in the following subsection.

The maximum achievable offset power occurs when the mobile device lies as far as possible from any one of the fingerprinted sites, i.e. at the midpoint of the square formed by the four sites closest to the mobile. By increasing the grid spacing, this maximum displacement will also increase. Hence,  $\|\Delta\mathbf{P}\|$  is proportional to the grid spacing. The non-centrality of the distribution is attributed to the

offset term,  $\Delta\mathbf{P}$ , which shifts the peak of the pdf to  $\rho = \|\Delta\mathbf{P}\|$ ; and since  $f_{\rho_c}(0) = 0$  invariably, when the curve is shifted to the left, it also spreads out. Therefore, larger grid spacing also leads to more uncertainty in the pdf.

### 4.1.2 A Stochastic Model for the Correct Localization

The model in the previous subsection assumes that the location system associates the mobile's location to the site which has the smallest similarity metric. In this subsection, the same assumption is made. Under this assumption, the mobile device is correctly localized if the shadowing component of the measured power does not cause it to deviate closer to the signature power of a different site. In the sequel, a model for the probability of correct localization is developed.

#### 4.1.2.1 Model Description

Formally stated, the system correctly localizes the mobile device if the measured power,  $\hat{P}$ , is more similar to the fingerprinted power of site  $c$ ,  $P_c$ , than to the fingerprinted power of any another site  $i$ . The marginal probability of correct localization when considering a single site  $i$  can be expressed as

$$p(\rho_c \leq \rho_i) = p\left(\sum_{j=1}^{n_B} (\hat{P}_j - P_{cj})^2 \leq \sum_{j=1}^{n_B} (\hat{P}_j - P_{ij})^2\right). \quad (4.4)$$

By expanding and collecting terms, the expression can be reduced to

$$p(C_i \leq 0), \quad (4.5)$$

where  $C_i = \sum_{j=1}^{n_B} 2\hat{P}_j\beta_{ij} + \Delta P_{ij}$  is a newly defined random variable with associated constants  $\Delta P_{ij} = (P_{ij} - P_{cj})$  and  $\beta_{ij} = (P_{cj}^2 - P_{ij}^2)$ . Note that vector  $\Delta\mathbf{P}_i$  is the offset power between the fingerprints of sites  $i$  and  $c$ . The vector  $\beta_i$  is a second-order offset. It follows that since  $\hat{P}_j$  is normally distributed due to shadowing,  $C_i$  is also normally distributed, however with mean and variance

$$\begin{aligned} \mu_{C_i} &= \sum_{j=1}^{n_B} 2P_{ij}\beta_{ij} + \Delta P_{ij} \\ \sigma_{C_i}^2 &= \sum_{j=1}^{n_B} (2\beta_{ij}\sigma)^2. \end{aligned} \quad (4.6)$$

Now, the total probability of correctly localizing the mobile device to site  $c$ —total here implies when considering all of the other  $n_M - 1$  sites, not just site

$i$ —can be computed. This probability,  $p(C)$ , can be expressed as the joint probability of all the other sites having a greater similarity metric than site  $c$ :

$$p(C) = p(C_1 \leq 0, C_2 \leq 0, \dots, C_{c-1} \leq 0, C_{c+1} \leq 0, \dots, C_{n_M} \leq 0). \quad (4.7)$$

Note, however, that the  $n_M - 1$  events above are interdependent. This can be seen by considering a simple example with only one base station in the deployment area for which  $P_{c1} < \hat{P}_1 < P_{i1} < P_{i+1,1}$ . It follows that  $p(C_i \leq 0)$  implies  $p(C_{i+1} \leq 0)$ . Unfortunately, computing the joint probability in Eq. (4.7) results in a complicated expression. Rather, as an approximation, the events are considered to be independent. As such,

$$p(C) \approx \prod_{\substack{i=1 \\ i \neq c}}^{n_M} p(C_i \leq 0). \quad (4.8)$$

The validity of this approximation is examined in the paper. It shows that for  $n_B > 2$ , which is the case in most practical implementations, the approximation holds very well. This demonstrates that adding base stations to the system decorrelates the events. The events were further decorrelated because the experiments were conducted in non-line-of-sight conditions—conditions for which the size of the random component (shadow fading) is yet larger. The details of the experiments are included next.

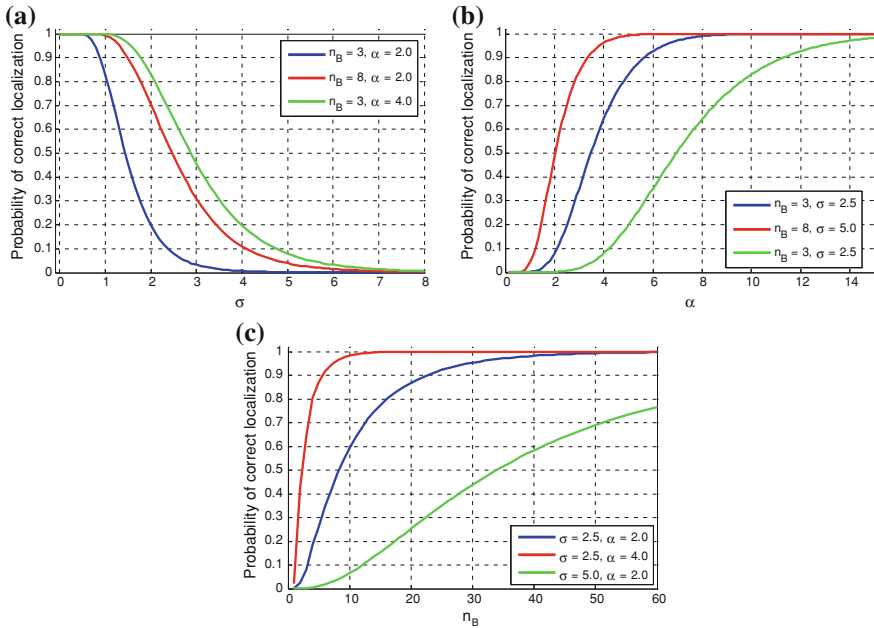
#### 4.1.2.2 Performance Evaluation

The probability of correctly localizing the mobile device—the performance metric of the system—was analyzed by considering an example deployment with 25 sites arranged as a  $5 \times 5$  grid (see Fig. 4.2). The grid spacing was 1 m. The  $n_B$  base stations were positioned randomly at grid points around the outermost square and the mobile device was fixed at the grid center. The simple path loss model from (2.24) was employed. In it, the path loss is modeled as a deterministic function of the distance,  $d_{ij}$ , between site  $i$  and base station  $j$ :

$$L(d_{ij}) = L^0 + 10\alpha \log_{10} \left( \frac{d_{ij}}{d^0} \right). \quad (4.9)$$

The reference path loss and the reference distance were set respectively to  $L^0 = 37.7$  dB and  $d^0 = 1$  m and the base stations all transmitted at  $P^{TX} = 15$  dBm. Neglecting the transmitter and receiver antenna gains, the received power is computed as the transmit power minus the path loss, plus a shadowing component, or:

$$P_{ij} = P^{TX} - L(d_{ij}) + S. \quad (4.10)$$



**Fig. 4.4** The probability of correct localization. **a** Probability of correct localization as a function of the standard deviation of shadow fading for different numbers of base stations and path loss exponents. **b** Probability of correct localization as a function of path loss exponent for different numbers of base stations and shadow fading standard deviations. **c** Probability of correct localization as a function of the number of base stations for different shadow fading standard deviations and path loss exponents

The shadowing was assumed to be distributed as a zero-mean Gaussian with standard deviation  $\sigma$ . In the fingerprinting stage, the feature vector of a site was calculated as the deterministic received power—i.e., with zero shadow fading—from each base station. Only  $n_M = 9$  of the sites were fingerprinted: the center grid point and its eight adjacent points.

Figure 4.4a investigates the effect of the shadowing parameter  $\sigma$  on the probability of correct localization. The set of parameters associated with the green curve could represent the outdoor propagation environment ( $\alpha = 4$  is a typical value for the path loss exponent) with three base stations. For these parameters, the performance of the system is shown to degrade rapidly—from a probability above 0.8 to a value below 0.2—as the standard deviation increases from 2 dB beyond 4 dB. But typical values of  $\sigma$  outdoors can be as high as 10 dB in urban environments; even indoors, the experiments in Gentile et al. (2008) report values in the range 2.8–5.4 dB. This demonstrates that a localization resolution of 1 m is practically unattainable for these parameters—which correspond to the best-case scenario of the three shown—even with such fine grid spacing, which in practice would require a laborious fingerprinting stage. In fact, most physical implementations of memoryless fingerprinting systems using RSS report errors above 2 m (Bahl and



Padmanadhan 2000; Brunato and Battiti 2005). As we shall see in later sections, implementing memory systems or using different features, namely the channel impulse response to better characterize the sites, can greatly improve results.

Figure 4.4b investigates the effect of the path loss exponent on the system. The performance is shown to improve as  $\alpha$  increases, meaning that the system localizes better in a harsher propagation environment. The explanation for this is that the fingerprinted power between sites dropped off more rapidly, enhancing the discriminatory properties of the system, thereby decreasing the chances of misclassification. Hence, the fingerprinting technique exploits the very weaknesses of the RSS ranging technique—which is intended for quasi-line-of-sight conditions only—described in Chap. 2. So, in fact, the two techniques are complementary.

The probability of correct localization—that is selecting the correct nearest neighbor grid point—improves as the grid spacing increases. The effect seen through the model is equivalent to increasing the path loss exponent. This is because, since the sites are farther apart, there is more variability in the signal strength between them. In fact, when plotting the performance metric as a function of grid spacing, the curves look very similar to the ones in Fig. 4.4b. Of course, the disadvantage of larger grid spacing is that it lowers the maximum attainable resolution.

Finally, the number of base stations in the system was varied. As seen in Fig. 4.4c, the performance of the system for the parameters corresponding to the red curve, which assumes harsh propagation environment ( $\alpha = 4$ ) and low shadow fading standard deviation ( $\sigma = 2.5$ ), stabilizes at  $n_B = 5$ . Beyond this value it continues to increase, but at a diminishing rate. In a more favorable propagation environment, there is a bit more benefit from adding base stations (blue curve), and with higher shadow fading (green curve) there is more consistent benefit. As mentioned earlier, the effect of shadowing weighs more heavily on the system as the number of base stations increases; this is seen by the shallower slope of the green curve with respect to the blue. Yet, this effect is offset by the benefit of greater identifiability; hence the performance continues to improve monotonically.

## 4.2 Memoryless Systems

The analysis in the previous section assumes the nearest neighbor method—the most simplistic of mapping algorithms—is employed to determine the mobile's location from a measured feature. However, more sophisticated methods, such as the  $k$ -nearest neighbor method ( $k$ NN), probabilistic methods, neural networks, support vector machines, and the smallest  $M$ -vertex polygon method in Liu et al. (2007) can enhance localization accuracy. For example, Agiwal et al. (2004) introduced the LOCATOR algorithm, which is an RSS-based fingerprinting technique incorporating a number of different approaches. Specifically, in the fingerprinting stage, the radio map is subdivided into clusters to reduce the computational cost in the localization stage. The authors further use RSS distribution functions at the sites and interpolations to improve performance. In Moustafa and Ashok (2005),

the Horus RSS-based system models the RSS distribution received from base stations through parametric and non-parametric distributions, exploiting this information to reduce temporal variations in the radio map. Also, Fang et al. (2008) demonstrated further improvements by using an RSS averaging technique on a logarithmic scale to mitigate noise resulting from multipath.

The purpose of this section is to provide an overview of some of the aforementioned methods. Specifically, we present the comparison which was published in Brunato and Battiti (2005) between the weighted  $k$ -nearest neighbor method, support vector machines, Bayesian inference, and neural networks. As mentioned earlier, although these methods implement the received-signal-strength feature, they can be readily extended to features such as the channel impulse response or the frequency channel coherence function.

### 4.2.1 The Weighted $k$ -Nearest Neighbors Method

The first of the mapping algorithms we investigate is the  $k$ -Nearest Neighbor method, which is just an extension of the nearest neighbor method providing enhanced robustness to shadowing. Precisely, rather than map the mobile's location to the single nearest neighbor site, the  $k$  nearest neighbor sites are employed, where  $k$  is a fixed constant. In practice, the mobile's location is estimated as the centroid of the  $k$  site locations—together these sites form subset  $\mathbf{K}$ —which have the smallest similarity metrics among all the sites. A refinement of the method is the weighted  $k$ NN method proposed in Brunato and Battiti (2005), which scales the contribution of each by the reciprocal of the similarity metric. Specifically, the mobile's location is estimated as a linear combination from the subset:

$$\tilde{\mathbf{x}} = \frac{\sum_{i \in \mathbf{K}} \frac{\mathbf{x}_i}{\rho_i + \rho_0}}{\sum_{i \in \mathbf{K}} \frac{1}{\rho_i + \rho_0}}. \quad (4.11)$$

As such, the location will fall within the convex hull of the site locations. By associating to location  $\mathbf{x}_i$  a weight inversely proportional to the similarity metric  $\rho_i$ , greater importance is given to sites whose signature power is closer to the measured power. The constant  $\rho_0$  is a small quantity added to ensure numerical stability when the similarity metric is close to zero, and the denominator of (4.11) serves to normalize the weights such that their sum is equal to one.

### 4.2.2 Support Vector Machines

Support vector machines (SVM) were developed in the area of supervised machine learning in order to solve nonlinear regression and statistical classification

problems. RSS-based fingerprinting methods based on support vector machines have been reported in Wu et al. (2004), Li Wu et al. (2007). In Brunato and Battiti (2005), they provide a direct mapping from the measured power at the mobile device to its estimated location through nonlinear regression<sup>1</sup>—nonlinear regression on the training pairs  $(\mathbf{x}_i, \mathbf{P}_i)$ ,  $i = 1 \dots n_M$ . Two mappings from the measured power vector,  $\hat{\mathbf{P}}$ , to the estimated  $(x, y)$ -coordinates of the mobile location,  $\mathbf{x}$ —denoted as  $x(\hat{\mathbf{P}})$  and  $y(\hat{\mathbf{P}})$ —are generated separately. Henceforth, we concentrate on the  $x$ -mapping, as the method applies equivalently to the  $y$ -mapping.

The  $x$  mapping can be expressed as a weighted sum of  $M$  prescribed nonlinear functions,  $g_m(\cdot)$ ,  $m = 1 \dots M$ , or

$$x(\hat{\mathbf{P}}) = \sum_{m=1}^M w_m g_m(\hat{\mathbf{P}} - \bar{\mathbf{P}}) + \bar{x}, \quad (4.12)$$

where  $\bar{x} = \sum_{i=1}^{n_M} x_i$  is the  $x$ -centroid and  $\bar{\mathbf{P}} = \sum_{i=1}^{n_M} \mathbf{P}_i$  is the mean power vector. The solution to the regression yields values for the weights  $w_m$ . For instance, if  $g_m(\cdot) = (\cdot)^{m-1}$  is selected,  $x(\hat{\mathbf{P}})$  is represented by an  $(M - 1)$ -th-degree polynomial, where the weights form the associated set of coefficients.

The regression is obtained by solving a convex quadratic program with the following objective function:

$$C \sum_{i=1}^{n_M} \zeta_i + \frac{1}{2} \sum_{m=1}^M w_m^2. \quad (4.13)$$

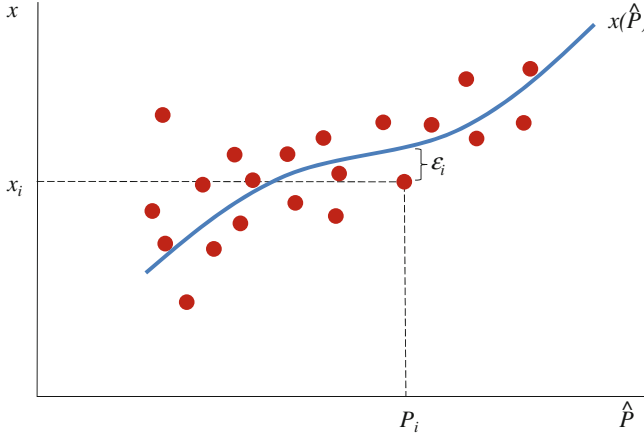
The main objective of the program, which is embodied by the first term, is to find the mapping which yields the best fit to the training pairs; this is achieved by minimizing the sum of residuals  $\zeta_i = |x_i - x(\mathbf{P}_i)|$  over all the pairs. The secondary objective, embodied by the second term, is to reduce the complexity of the mapping such that it can be represented in the lowest dimensional space, where  $M$  is the maximum dimension; this is achieved by minimizing the norm of the weights. The constant,  $C$ , balances the importance of the two objectives. Then the quadratic program can be stated completely as:

$$\begin{aligned} \min \quad & C \sum_{i=1}^{n_M} \zeta_i^+ + \zeta_i^- + \frac{1}{2} \sum_{m=1}^M w_m^2 \\ \text{subject to} \quad & \begin{cases} x_i - x(\mathbf{P}_i) \leq \zeta_i^+ \\ x(\mathbf{P}_i) - x_i \leq \zeta_i^- \\ \zeta_i^+, \zeta_i^- \geq 0 \end{cases} \end{aligned} \quad (4.14)$$

By decomposing the residuals into positive and negative components, as  $\zeta_i = \zeta_i^+ - \zeta_i^-$ , the absolute values on the residuals are removed such that the

---

<sup>1</sup> Their application to statistical classification is similar.



**Fig. 4.5** The Support Vector Machine (SVM) mapping between the measured received signal power space,  $\hat{P}$ , and the mobile location space,  $x$

problem can be written in standard form. Figure 4.5 illustrates an example regression for the function  $g(P) = P^3$  in the one-dimensional power vector space.

As is often the case in convex programming, here it is more practical to solve the dual quadratic program in (4.15) instead by introducing Lagrange multipliers,  $(\lambda_i^+, \lambda_i^-)$ ,  $i = 1 \dots n_M$  (Smola and Schoelkopf 2004):

$$\begin{aligned} \max \quad & -\frac{1}{2} \sum_{i=1}^{n_M} \sum_{k=1}^{n_M} (\lambda_i^+ - \lambda_i^-) K(\mathbf{P}_i, \mathbf{P}_k) (\lambda_k^+ - \lambda_k^-) + \sum_{i=1}^{n_M} x_i (\lambda_i^+ - \lambda_i^-) \\ \text{subject to} \quad & \sum_{i=1}^{n_M} (\lambda_i^+ - \lambda_i^-) = 0, \\ & 0 \leq \lambda_i^+, \lambda_i^- \leq C \end{aligned} \quad (4.15)$$

where  $K(\mathbf{P}_i, \mathbf{P}_k) = \sum_{m=1}^M g_m(\mathbf{P}_i - \mathbf{P}) g_m(\mathbf{P}_k - \mathbf{P})$  is known as the kernel function. The solution to the dual problem yields the values for  $(\lambda_i^+, \lambda_i^-)$ . From them, the components of the weight vector in (4.12) can be found as

$$w_m = \sum_{i=1}^{n_M} (\lambda_i^+ - \lambda_i^-) g_m(\mathbf{P}_i - \mathbf{P}). \quad (4.16)$$

### 4.2.3 Neural Networks

In contrast to the well-defined mathematical formulation provided by Support Vector Machines, a “black box” approach for generating the mapping between the measured power vector space and the estimated mobile location space is through

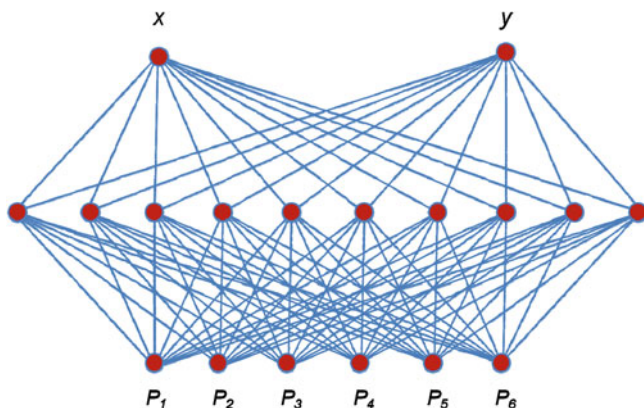
the application of neural networks. Neural networks are powerful tools for solving ill-posed problems, i.e., problems for which the causes of certain observations are either not understood or too complex to define mathematically, and for which there is often no unique solution. This applies to fingerprinting systems in that an observed radio frequency feature value—depending on the number of base stations in the system—may correspond to multiple locations in the survey area; moreover, the value depends on the structure of the environment (building blueprint, wall materials, furniture characteristics, etc.). For such systems, there is no attempt to explicitly model the complex propagation environment which causes these observations. Rather, it is simply acknowledged that such a nonlinear relationship exists. By observing the RF signatures at specific locations, the neural network learns the relationship. Neural network methods for RSS-based location fingerprinting have been reported in Battiti et al. (2002), Edgar et al. (2004), Brunato and Battiti (2005).

A neural network is a network composed from entities, known as neurons, which have multiple input ports and a single output port. In Brunato and Battiti (2005), the multilayer perceptron neural network is implemented. The multilayer perceptron, in particular, is a feedforward network partitioned into distinct layers. Feedforward means that the input of a neuron in one layer is connected only from the outputs of a neuron in the immediate lower layer. Each connection has an associated weight which serves to scale the output value between the two layers. In the RSS-fingerprinting application, the inputs to the lowest layer of the network are the  $n_B$  elements of the measured power vector—there is one neuron for each base station. Likewise, the outputs of the highest layer are the two coordinates of the location vector—there is one neuron for each coordinate dimension. Figure 4.6 shows a diagram of the network.

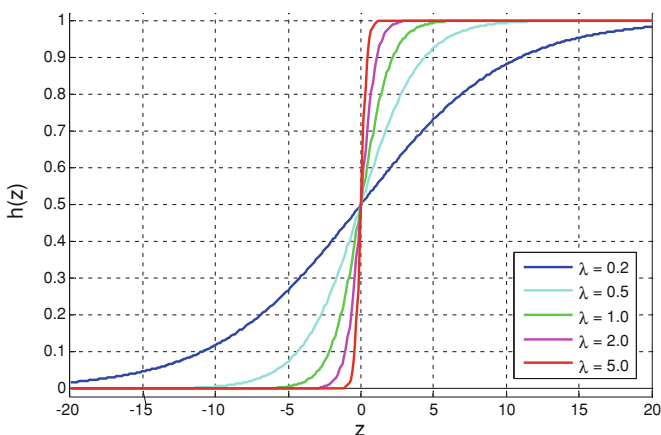
The role of the individual neuron in the network is simply to compute its output value from the collection of its inputs. This is executed by summing over the input values and then mapping the sum to the output through an activation function. By choosing a nonlinear function for the neuron, the network is capable of representing any nonlinear function for the network as a whole. In fact, the Cybenko theorem, also known as the universal approximation theorem, states that a feedforward network with a single hidden layer and a finite number of neurons can approximate any continuous function (assuming a “well-behaved” activation function) (Cybenko 1989). A commonly used activation function for the perceptron is the sigmoid function:

$$h(z) = \frac{1}{1 + e^{-\lambda z}}, \quad (4.17)$$

where  $\lambda$  controls the linearity of the function. Figure 4.7 illustrates the curve for several values of the parameter. As in the SVM framework, neural networks can be implemented for both regression problems and classification problems. For small values of  $\lambda$ , the function is linear around  $z = 0$  and then saturates at the extrema. This range of  $\lambda$  is applicable to regression problems, such as ours, so that the



**Fig. 4.6** A three-layer feed forward neural network. The inputs to the first layer are the measured received signal strengths from six base stations. The outputs are the two-dimensional coordinates of the estimated location of the mobile device. The network has one hidden layer



**Fig. 4.7** The sigmoid function

output can assume continuous values. As  $\lambda$  approaches infinity,  $h(z)$  becomes a step function. This range is applicable to classification problems for which the output is either a one or a zero, meaning that the input either belongs to a certain class or belongs respectively to a different class.

The network learns the mapping through an iterative algorithm, such as the well-known Backpropagation Algorithm, which tunes the connection weights according to the input/output excitations. During each iteration, by clamping the inputs and outputs of the network with the values of the training pair  $(\mathbf{x}_i, \mathbf{P}_i)$ , the weights are adjusted such that the network yields  $\mathbf{x}_i$  as an output given  $\mathbf{P}_i$  as an input. The details of network design and training can be found in Bose (1995).

#### 4.2.4 Bayesian Inference

Probabilistic approaches for estimating the location of a mobile device in fingerprinting systems have been reported in Roos et al. (2002), Youssef et al. (2003), Fox et al. (2003), Madigan et al. (2005), Kushki et al. (2007). As in the  $k$ NN, SVM, and neural network frameworks, the estimated location is not constrained to any one of the discrete fingerprinted sites, meaning that it can assume any position throughout the deployment area. The problem is posed in the framework of Bayesian inference: given the measured power vector,  $\hat{\mathbf{P}}$ , the posterior probability (or simply the posterior),  $p(\mathbf{x}|\hat{\mathbf{P}})$ , that the mobile is located at position  $\mathbf{x}$  is calculated for all candidate positions in the area. Then, from this probability, the mobile's location is estimated either through Maximum Likelihood as

$$\tilde{\mathbf{x}} = \max_x \arg p(\mathbf{x}|\hat{\mathbf{P}}) \quad (4.18)$$

or as an expected value over the area:

$$\tilde{\mathbf{x}} = \int_x \mathbf{x} \cdot p(\mathbf{x}|\hat{\mathbf{P}}) d\mathbf{x}. \quad (4.19)$$

The posterior probability,  $p(\mathbf{x}|\hat{\mathbf{P}})$ , can be viewed as a mapping from the power vector space to the mobile location space. In the SVM and neural network frameworks, such mappings are computed through some sort of nonlinear regression on the training pairs. In the Bayesian framework, however, a mapping is first computed in the opposite direction, i.e., from  $\mathbf{x}$  to  $\hat{\mathbf{P}}$ . This inverse mapping, denoted as  $p(\hat{\mathbf{P}}|\mathbf{x})$ , is known as the likelihood function (or simply the likelihood) and effectively serves as the RSS signature for the site. It is the probability function that the power vector  $\hat{\mathbf{P}}$  will be measured if the mobile device is at  $\mathbf{x}$ . The benefit of this approach is that likelihood can be computed directly from the training pairs. For example, in (Roos), (Kuschki), (Fox), (Madigan) the likelihood function is constructed from the histogram of RSS values registered at each site. (More details about how to generate the histogram are provided in Sect. 4.3). Once the likelihood is computed, it is related back to the posterior probability through Bayes' Rule, as we shall see in the sequel.

As an alternative to constructing histograms at the discrete sites, (Brunato and Battiti 2005) invoke a path loss model to calculate the likelihood function. The path loss model enables generating RSS signature values at continuous points throughout the survey area—rather than at discrete points only—in the hope of improving localization resolution. The path loss model employed is similar to the traditional model in (4.9), however, it also accounts for the attenuation of the walls between a base and mobile pair. This more comprehensive path loss model can be expressed as:

$$L_j(\mathbf{x}) = L_j^0 + 10\alpha_j \log_{10} \left( \frac{d_j(\mathbf{x})}{d^0} \right) + L_j^w \cdot n_j^w(\mathbf{x}) \quad (4.20)$$

Note that, in order to represent the radio environment more precisely, each base station  $j$  has its own path loss,  $L_j(\mathbf{x})$ . The first term is the associated reference path loss,  $L_j^0$ , and the second term is the propagation loss, where  $\alpha_j$  is the loss exponent and  $d_j(\mathbf{x})$  is the distance between the base station and the mobile device. The last term is the penetration loss due to walls, with  $n_j^w(\mathbf{x})$  denoting the number of walls between the base and the mobile and  $L_j^w$  denoting the penetration loss per wall.

The unknown parameters of the path loss model can be extracted through the data points given by the training pairs  $(\mathbf{x}_i, \mathbf{P}_i)$ ,  $i = 1 \dots n_M$ . To this end, recall from Eq. (4.10) that the deterministic received power at the mobile is given from the loss as

$$P_j(\mathbf{x}) = P^{TX} - L_j(\mathbf{x}), \quad (4.21)$$

where  $P^{TX}$  is the known transmit power. Then for each base station  $j$ , the training pair  $(\mathbf{x}_i, \mathbf{P}_i)$  furnishes exactly one linear equation with three unknowns from (4.21). The system of  $n_M$  equations, which is overdetermined for  $n_M > 3$ , can be solved for the values of  $(L_j^0, \alpha_j, L_j^w)$  through Least Squares Regression. Note that it is also possible to assume the same loss model for all base stations by removing the index  $j$  in (4.20), however the authors report that this causes degradation in performance.

With the parameters of the path loss model in hand, the likelihood function can now be obtained. Recall that in the shadow fading model, the measured power,  $\hat{P}_j$ , from base station  $j$  at location  $\mathbf{x}$  deviates from the deterministic power,  $P_j(\mathbf{x})$ , by the random variable  $S$ . In other words,

$$\hat{P}_j = P_j(\mathbf{x}) + S \quad (4.22)$$

Since  $S$  is a zero-mean normally distributed random variable, the likelihood that  $\hat{P}_j$  was measured at location  $\mathbf{x}$  is given through the Gaussian kernel<sup>2</sup>:

$$p(\hat{P}_j|\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\hat{P}_j - P_j(\mathbf{x}))^2}{2\sigma^2}}. \quad (4.23)$$

For simplicity of computation—although not always true in practice—the signals from the  $n_B$  base stations are assumed to experience independent and identically distributed shadowing such that the received powers from each base station are also statistically independent. Note that this is directly related to the independence assumption in (4.8), which was experimentally shown to be a valid approximation for  $n_B > 2$ . This assumption is also made in (Roos), (Kuschki), (Fox), (Madigan). As a result, the likelihood of the measured power vector can be calculated as the product of the measured powers from each of the base stations:

---

<sup>2</sup> On a similar note, in (Roos), (Kuschki), (Fox), (Madigan) mathematical expressions which are close to the Nadaraya-Watson Kernel regression are developed.



$$p(\hat{\mathbf{P}}|\mathbf{x}) = \prod_{j=1}^{n_B} p(\hat{P}_j|\mathbf{x}) \quad (4.24)$$

Finally, the likelihood,  $p(\hat{\mathbf{P}}|\mathbf{x})$ , is related back to the posterior probability,  $p(\mathbf{x}|\hat{\mathbf{P}})$ , through Bayes' Rule:

$$p(\mathbf{x}|\hat{\mathbf{P}}) = \frac{p(\hat{\mathbf{P}}|\mathbf{x}) \cdot p(\mathbf{x})}{p(\hat{\mathbf{P}})} \quad (4.25)$$

If all the locations throughout the survey area are visited with equal frequency, the prior probability or simply the prior,  $p(\mathbf{x})$ , is uniformly distributed. Otherwise, if certain locations have higher or lower frequencies, the prior will be distributed proportionately; as a simple example, in most households more time is spent in the living room than in the attic. The value of  $p(\hat{\mathbf{P}})$  is computed through the law of total probability:

$$p(\hat{\mathbf{P}}) = \int_{\mathbf{x}} p(\hat{\mathbf{P}}|\mathbf{x}) \cdot p(\mathbf{x}) d\mathbf{x} \quad (4.26)$$

### 4.2.5 Comparison of Methods

The specific parameters implemented for comparing the four methods described in this section are provided in Brunato and Battiti (2005). The test experiments were conducted in a deployment area of roughly 750 m<sup>2</sup>. The area was partitioned into five rooms and in each room a separate Wi-Fi base station was deployed. While for the most part LOS conditions existed within the individual rooms, the walls throughout the area between the base stations and the mobile device created NLOS conditions. The fingerprinted sites were spaced at about 3.5 m apart, for a total of 257 sites in the area. For the parameter settings in the paper, the weighted  $k$ -nearest neighbor and the support vector machine methods delivered the best performance, both averaging a location error of about 3 m. While the computational complexity for training the  $k$ NN is lower than that of the SVM, the latter boasts a much lower complexity in the localization stage. The average location error for the neural network method was about 3.2 m, but the time required for tuning the 60 weights was the highest among all methods; once tuned, however, the neural network localized the quickest. The Bayesian interference was both computationally inefficient and also sustained the worst average error of 3.35 m. The authors attributed the poor performance to the adopted path loss model in (4.20) with a total of only 20 tunable parameters—four for each of the five base stations. The Bayesian method required only a few training points for parameter fitting but, once fit, providing more training points did not improve the results further. On the other hand, the neural network, with a total of 60 tunable weights, offered a better degree of fitting.

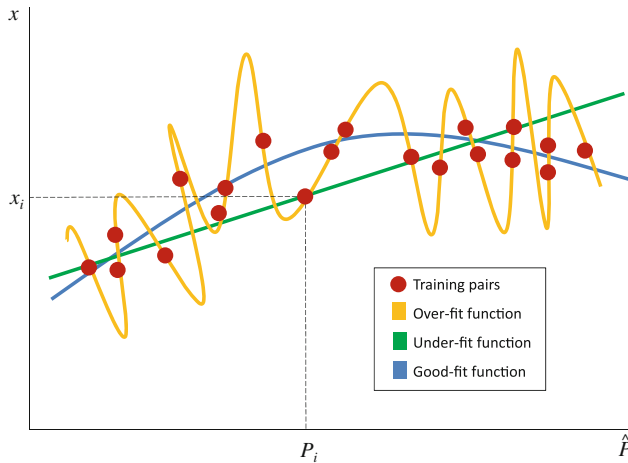
Another performance evaluation of different RSS-based fingerprinting methods is presented in Lin and Lin (2005). The authors compare the  $k$ NN, a probabilistic method, and neural networks. The results of the analysis and experiments reveal that the  $k$ NN reports the best overall performance for indoor positioning. The performance of histogram, nearest neighbor, parametric and kernel location fingerprinting methods were evaluated in Honkavirta October (2008), Honkavirta et al. March (2009). Again, the results revealed that the  $k$ -nearest neighbor method fared the same or better than the other methods depending on the environment.

In practice, it is difficult to rank the performance of the four methods described here because they are sensitive—each to one extent or another—to the choice of implementation parameters. As a matter of fact, the methods are more similar to each other than they are different—all essentially just fit a curve to the training pairs. The parameters selected for each determine to what degree the fitting can be achieved. In support vector machines, the degree of freedom increases with the order of the nonlinear functions  $g_m(\cdot)$ ,  $m = 1 \dots M$  in (4.12) and with the number  $M$  of functions itself. In addition—and more explicitly—by increasing the parameter  $C$  in Eq. (4.13), greater importance is given to the minimizing the fitting error; in contrast, by decreasing this parameter, the system order is minimized. In neural networks, the relationship to the system order is even more explicit: increasing the number of neurons in the hidden layer increases the degree of freedom. In Bayesian inference, as just mentioned, the degree of freedom is dictated by the number of unknowns in the path loss equation in Eq. (4.20). Finally, in the  $k$ -nearest neighbor method the estimated location is determined as a curve interpolated between the  $k$ -nearest neighbors. By increasing the value of  $k$ , although more robust to measurement error, the estimated location is more constrained.

For illustrative purposes only, Fig. 4.8 shows three curves fit to a set of training pairs (red). The orange curve represents a function which is overfit; in order to reduce the fitting error to zero, a high-order curve is allowed. While the error is zero for the set of training pairs, the curve does not interpolate well between the training pairs. The large oscillations indicate that a small change in the measured RSS vector maps to a completely different location, making for an unstable system. On the other hand, the green curve represents an underfit function; because the function has only a few degrees of freedom, it is very robust to fluctuations in signal strength. At the same time, the poor fitting to the training pairs can also lead to large location errors. Lastly, the blue curve presents a good balance between location accuracy and robustness.

### 4.3 Memory Systems

Thus so far we have considered only memoryless systems, which estimate the location of a mobile device based solely on the received signal strength observed at a single instant in time. While these systems may deliver acceptable performance



**Fig. 4.8** The four methods presented in this section—each through a different algorithm—generate some mapping between the signal strength space and the location space. The parameters of each determine the degree of fitting to the training pairs. Shown here are three fits for illustrative purposes

for some applications, by integrating observations available from previous time instants as well, both precision and stability can be enhanced. In this section, we describe techniques first developed to solve the *wake-up robot problem* (Burgard et al. 1996) which have been adapted to fingerprinting. The scope of wake-up robot problem is for a robot, which is placed in an arbitrary environment, to discern its position by gathering and processing sensory data with no prior knowledge. Chapter 9 is completely dedicated to these techniques—often referred to as Simultaneous Localization and Mapping (SLAM)—with specific application to inertial based localization. In the following, we first investigate a technique which is an adaptation of the Bayesian interference method introduced in Sect. 4.2.4 to memory systems. We then present an evolution of this technique, known as grid-based Markov localization, which delivers enhanced stability.

### 4.3.1 Bayesian Inference in Memory Systems

In this section, we consider an application for which the orientation of the mobile device, in addition to its location, is estimated. This is achieved by augmenting the fingerprinted information gathered at site  $i$ —previously only the location coordinates,  $x_i$ , were fingerprinted—with an orientation identifier denoted as  $\theta_i$ . The orientation identifier can assume one of two values:  $\theta_i = 1$  signifies that the user is facing a designated direction at the site while  $\theta_i = -1$  signifies that the user is facing the opposite direction. Of course more than just two orientations can be

incorporated, if desired. We now define a *state* variable  $s_k = \{\mathbf{x}_k, \theta_k\}$  for the mobile, which indicates both its location and its orientation. The mobile can lie in any of  $n_s$  possible states indexed through  $k = 1 \dots n_s$ . Note that for a total  $n_M$  fingerprinted sites with two orientations per site,  $n_s = 2n_M$ .

The Bayesian inference method enables constructing a time-varying posterior probability for the state of a mobile device. This probability, denoted as  $p(s_k | \hat{\mathbf{P}}^t, \dots, \hat{\mathbf{P}}^0)$ , represents the probability that the mobile lies in state  $k$  given the observations from initialization ( $t = 0$ ) to time  $t - 1$ . These observations are indexed accordingly as  $\hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0$ . The system is initialized by setting  $p(s_k | \hat{\mathbf{P}}^0) = 1/n_s, k = 1 \dots n_s$ . This means that in the absence of any observations, all locations are equally probable. Assuming the posterior at time  $t - 1$  has been computed, when the most recent observation,  $\hat{\mathbf{P}}^t$ , becomes available, Bayes' Rule is applied to compute the posterior at the next time step:

$$p(s_k | \hat{\mathbf{P}}^t, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) = \frac{p(\hat{\mathbf{P}}^t | s_k, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) \cdot p(s_k | \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0)}{p(\hat{\mathbf{P}}^t | \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0)} \quad (4.27)$$

The denominator in the equation above follows from the law of total probability as  $p(\hat{\mathbf{P}}^t | \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) = \sum_{k=1}^{n_s} p(\hat{\mathbf{P}}^t | s_k, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) \cdot p(s_k | \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0)$ , i.e. the sum of the  $k$ -indexed numerator over all the  $n_s$  states. The denominator effectively serves as a normalizing factor such that  $\sum_{k=1}^{n_s} p(s_k | \hat{\mathbf{P}}^t, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) = 1$ , meaning that the mobile device will lie necessarily in one of the  $n_s$  states at time  $t$ . The likelihood,  $p(\hat{\mathbf{P}}^t | s_k, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0)$ , in the numerator is the probability that the signal strength vector  $\hat{\mathbf{P}}^t$  will be observed when the mobile lies in state  $s_k$ . Since this probability is assumed to be stationary—meaning that the observed power when the mobile user is at a particular location and in a particular orientation is static over time—the readings from previous time instants have no bearing on it. This assumption can be stated mathematically as  $p(\hat{\mathbf{P}}^t | s_k, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) = p(\hat{\mathbf{P}}^t | s_k)$ .

The probability  $p(\hat{\mathbf{P}} | s_k)$  represents the distribution of the received signal strength vector,  $\hat{\mathbf{P}}$ , when the mobile is in state  $s_k$ . In this application, the distribution acts as the RSS signature for the corresponding location and orientation of the mobile. Indeed it has been shown in Gentile and Klein-Berndt (2004), Ladd et al. (2005) that the signature varies not only by location but also by orientation. While the distribution of the received power from a single base station is often assumed to be normal, in fact it is typically more complex and even multimodel. Rather, for a more accurate characterization of the RSS signature, the same authors propose generating a histogram of signal strength values empirically from a training set,  $\mathbf{P}_k^l, l = 1 \dots L$  of  $L$  readings gathered at the mobile over a fixed window of time during the fingerprinting stage. Let  $h_{kj}(\zeta)$  stand for the histogram of signal strength values collected from base station  $j$  when the mobile is in  $s_k$ . The histogram can be expressed mathematically as

$$h_{kj}(\zeta) = \frac{1}{L} \sum_{l=1}^L \delta(P_{kj}^l - \zeta), \quad (4.28)$$

where  $\delta$  is the Kronecker delta function and  $\zeta$  is an indicator variable which spans the range of all possible signal strength values. The range depends on the specifications of the equipment used.

When the most recent observation becomes available, the likelihood probability is computed as a product of the measured power mapped by the histogram, or:

$$p(\hat{\mathbf{P}}^t | s_k) = \prod_{j=1}^{n_B} h_{kj}(\hat{P}_j^t). \quad (4.29)$$

To improve the accuracy of the system, the implementation in Gentile et al. (2004) actually fingerprints the signal strengths of packets both to and from the base stations as two separate readings. Each site will then have two histograms per base station rather than one, doubling the factors in Eq. (4.29). The expression is based on the same assumption, as in Eq. (4.24), of independent RSS value between the  $n_B$  base stations. In reality, the histograms of different states will be correlated to some degree, however the independence assumption yields good results regardless.

As suggested in Ladd et al. (2005), the stability of the system can be enhanced through a simple post-processing step, where a modified posterior probability is generated at each update as

$$\tilde{p}(s_k | \hat{\mathbf{P}}^t, \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) = (p(s_k | \hat{\mathbf{P}}^t, \dots, \hat{\mathbf{P}}^0) + u_1) \cdot (p(s_k | \hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0) + u_2). \quad (4.30)$$

The modified posterior filters any spurious values which may appear as spikes in the system at a single time instant due to glitches or erroneous observations. The values of  $(u_1, u_2)$  are small constants which keep the modified probability from collapsing to zero. Then from the modified posterior, the estimated state of the mobile at time  $t$  is given through the Maximum Likelihood Estimation as

$$\tilde{s}_k = \arg \max_k \tilde{p}(s_k | \hat{\mathbf{P}}^t, \dots, \hat{\mathbf{P}}^0). \quad (4.31)$$

### 4.3.2 Grid-Based Markov Localization

By integrating observations over a period of time, Bayesian inference can deliver enhanced stability over static localization. However, the method is still susceptible to large fluctuations in received signal strength due to fading, even while the mobile remains in the same state. So that these fluctuations are not converted into random motion, a sort of temporal averaging mentioned above is incorporated by post-processing the estimated output; in addition, estimated locations that do not

support human motion, such as hops between mutually distant sites in the deployment area are filtered out. While filtering [e.g. Kalman filtering (Kalman 1960)] can improve location tracking in simple experiments—for instance, a mobile user walking up and down a hallway—it fails with more complex trajectories such as turning from a corridor into a room.

As an alternative to post-processing, in this subsection, the problem is cast in the framework of a Markov random process through which the dynamics of human movement can be encoded intrinsically as transition probabilities. In this framework, the system can be tuned for fluid motion and direction while still providing for abrupt changes where appropriate.

### 4.3.2.1 Motion Dynamics

In order to capture the system dynamics, the definition of a state  $s_k$ , which indicates a unique location and orientation of the user, is extended to a *sequence*  $s_k = \{s_k^1, \dots, s_k^n\}$ . A sequence is defined as a set of states ordered in time representing the last  $n$  states traversed by the mobile device, from time  $t - n + 1$  to time  $t$ . Accordingly,  $n_s$  now denotes the total number of possible sequences. Integrating more than a single state captures not only the location and orientation of the mobile at consecutive instants in time, but also the dynamics of the motion between the states. How these sequences are composed is discussed later in the subsection.

At time step  $t$ , the localization algorithm calculates the posterior probabilities of the sequences,  $p(s_k | \hat{\mathbf{P}}^0, \dots, \hat{\mathbf{P}}^t)$ , given the observations since initialization. A first-order Markov process governs the transition of the sequences from step  $t - 1$  to the next (Fox et al. 1999):

$$p(s_k | \hat{\mathbf{P}}^0, \dots, \hat{\mathbf{P}}^t) = \eta^t \cdot p(\hat{\mathbf{P}}^t | s_k) \sum_{\tilde{s}_k=1}^{n_s} p(s_k | s_{\tilde{k}}) \cdot p(s_{\tilde{k}} | \hat{\mathbf{P}}^0, \dots, \hat{\mathbf{P}}^{t-1}) \quad (4.32)$$

Note the similarity of the expression to Eq. (4.27). The only difference is the incorporation of the sequence transition probabilities,  $p(s_k | s_{\tilde{k}})$ : in the Bayesian framework, the posterior of  $s_k$  at  $t$  is computed only from the posterior of  $s_k$  at  $(t - 1)$ . In the Markov framework, rather, it is computed from the posteriors of all  $n_s$  sequences at  $t - 1$  through the sequence transition probabilities. The algorithm reports the output state of the system at each step  $t$  as  $\tilde{s}^n, \tilde{s} = \arg \max_k p(s_k | \hat{\mathbf{P}}^0, \dots, \hat{\mathbf{P}}^t)$ . Again, the normalization factor  $\eta^t = 1 / \sum_{k=1}^{n_s} p(s_k^t | \hat{\mathbf{P}}^0, \dots, \hat{\mathbf{P}}^t)$  enforces the law of total probability and, since an observation at time  $t$  affects only the state of a sequence corresponding to the same time instant  $s_k^n$ , the likelihood can be simplified to  $p(\hat{\mathbf{P}}^t | s_k) = p(\hat{\mathbf{P}}^t | s_k^n)$ . As such, the value of  $p(\hat{\mathbf{P}}^t | s_k^n)$  is given from Eq. (4.29).

We now turn our attention to computing the sequence transition probabilities as in Gentile et al. (2004). First of all, in order to ensure spatiotemporal consistency

between back-to-back sequences, when the mobile is in sequence  $s_{\bar{k}}$  the sequences  $s_k$  to which it can transition at the next time step are restricted. This is implemented by setting  $p(s_k|s_{\bar{k}}) = 0$  if  $s_k$  does not meet the condition  $s_k^{l-1} = s_{\bar{k}}^l, l = 2, \dots, n$ ; in other words,  $s_k$  must be a left-shift of  $s_{\bar{k}}$  with replacement of only the  $n^{\text{th}}$  state,  $s_k^n$ , with a new state. Then, the other sequences, the so-called *allowed* sequences, are assigned a nonzero transition probability. The probability is assigned in order to promote fluid motion—that is motion which follows a predictable trajectory, such as the mobile moving down a corridor at a fixed velocity or slowing to a stop. If the ordered states of a sequence reflect fluid motion, the sequence is assigned a high probability and vice versa. The fluidness is characterized through an  $(n - 1)$ -tap filter. The filter is employed to predict the most likely  $n^{\text{th}}$  location in the sequence from the trajectory of the first  $n - 1$  locations:

$$\hat{x} = \sum_{l=2}^n \alpha^l \cdot \mathbf{x}_k^l = \sum_{l=1}^{n-1} \alpha^l \cdot \mathbf{x}_k^l, \quad (4.33)$$

where  $\alpha^l$  are the filter coefficients. Other non-finite impulse response filters, in particular the popular Kalman filter, may be applied alternatively. A Gaussian kernel maps the difference—between the actual location of the  $n^{\text{th}}$  state,  $\mathbf{x}_k^n$ , and its predicted location,  $\hat{x}$ —to the sequence transition probability:

$$p(s_k|s_{\hat{k}}) = \frac{1}{\gamma\sqrt{2\pi}} e^{-\frac{1}{2\gamma^2}\|\mathbf{x}_k^n - \hat{x}\|^2}. \quad (4.34)$$

A small difference (high probability) indicates that the sequence conforms well to the motion dynamics represented by the filter and a large difference (low probability) the opposite. The parameter  $\gamma$  controls the degree of Gaussian rolloff. Reducing the value of  $\gamma$  makes the sequence filtering more selective.

Even by restricting the sequences which are allowed, the number  $n_s$  may still grow exponentially large with  $n$ . Hence grid-based Markov localization can suffer from computational overhead and/or overcommitment of the memory requirements for the sequence space. Both, indeed, can present significant issues for location devices which are often very compact in size. The CONDENSATION algorithm, which falls into the general class of particle filters, offers a solution. Essentially, rather than maintaining the posterior probability for each discrete sequence in the model, the algorithm maintains only an abridged set of the most likely  $n_c \ll n_s$  sequences, i.e. the ones with the relatively largest associated values of  $p(s_k|\hat{\mathbf{P}}^{t-1}, \dots, \hat{\mathbf{P}}^0)$ . At the next step, these posteriors are updated to time  $t$ , as normal. And again, only the  $n_c$  sequences which have the relatively largest updated values are retained. The CONDENSATION algorithm has proven to be a powerful tool in recent years in the context of Bayesian estimation and computer vision. The details of the algorithm can be found in Isard and Blake (1998).

### 4.3.2.2 Motion Constraints

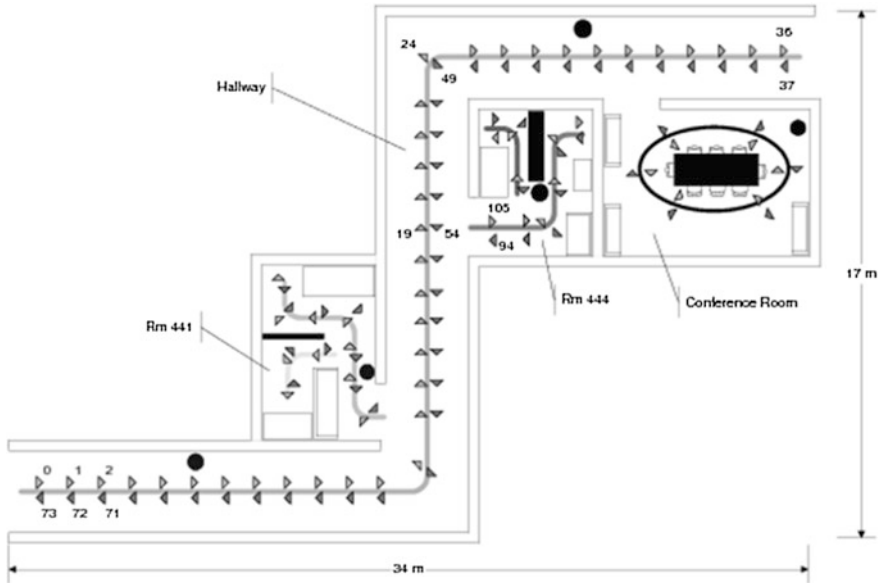
Certain applications, such as in emergency response, require precision location discrimination—knowing whether a firefighter lies in a particular room or in the one adjacent to it can make all the difference in saving a life. While the received signal strengths from multiple base stations alone may not suffice to make this distinction—depending on the material and thickness of a wall, the RF signature may differ little on its opposite sides—tracking the path of the mobile user as he or she enters the room may. As we shall see, in the framework of Markov localization, mobile user tracking can be realized by restricting the allowed sequences further by applying motion constraints.

The percentage of walking space in a typical office environment, which is furnished with desks, bookshelves, cubicles, and other furniture and equipment, ranges between 25 and 40 % of the total deployment area. The same is true in many other environments, namely in residential environments and in public environments such as libraries, supermarkets, etc. The presence of these obstacles severely constrains the paths along which humans can move about. Figure 4.9 illustrates a typical office environment. Six paths, displayed in different shadings of gray, connect any two fingerprinted sites in the environment. The pair of numbered arrows represents the two states corresponding to the opposite orientations at each site, splitting each path into two *tracks*. By fingerprinting each site with the antenna orientation aligned with the heading of the person, a mobile user walking forward on a path follows the states on either one track or on its complementary track. Under the assumption that a human walks only forward and that the antenna orientation remains constant with respect to the person's heading, motion constraints can be imposed such that the mobile can be localized as moving only along the tracks.

Motion constraints are applied to the Markov model such that a state can transition only to a spatially adjacent state from one time instant to the next. Consequently, the mobile must traverse a sequence of adjacent states or *neighbors* in order to reach any one state in the model from another. This is implemented by assigning the appropriate sequence transition probabilities a zero value. Recall that the same was explained earlier in application to restricted sequences. This mechanism, which allows only those sequences in the model which conform to the motion constraints (and restricts those which do not), turns out to be a highly effective manner to reconstruct a path from a series of observations during the localization stage. Classical Kalman filtering may predict the trajectory of a human advancing through a wall because it considers only the locations on the trajectory; motion constraints, rather, provide a blueprint of the area encoded through the sequence transition probabilities. The desired effect is that the system realizes that humans must go through doors in order to reach locations on opposite sides of a wall.

We now turn to the description of how neighbors and tracks are encoded in the Markov model. Most states have three neighbors: (1) itself—to allow stationary motion in time; (2) the next state on the same track—to allow motion in the same

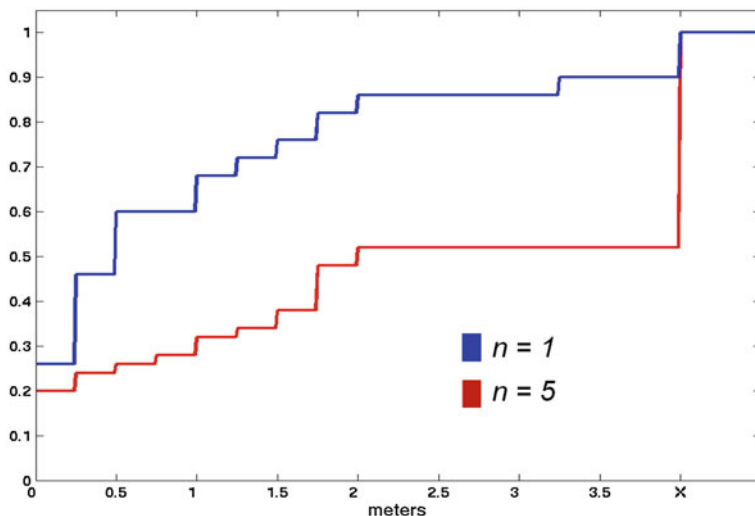




**Fig. 4.9** A typical office environment. Shown here are 124 numbered states divided into six paths—each path is shaded differently. The five base stations in the survey area are labeled as solid circles. Each path is split into tracks which allow motion in opposite directions, as indicated by the arrows

direction; (3) the state at the same location on the opposite track—to allow a change in direction. Exceptions occur for states at the end of tracks with no next state; they have only two neighbors. Another exception is for states falling at T-junctions or crossroads between two paths; additional neighbors enable the mobile to switch paths. In order to promote motion along tracks, sequences which contain more than one track transition are restricted. Moving backwards on a track is also not permitted in this particular implementation; such motion, however, is actually common in some applications; for example, firefighters walk backwards pulling hoses and crawl backwards downstairs. Of course the system can be tuned accordingly. Also, in large, open areas, a grid of states, rather than tracks, can be created and the appropriate motion constraints applied.

As an experiment, in Gentile et al. (2004) a system with sequence length  $n = 5$  was tested against a benchmark system with length  $n = 1$  in the office environment depicted in Fig. 4.9. For each trial, the localization error was recorded either as (1), the distance between the estimated location and the ground-truth location; or as (2), a logical error  $X$  when the mobile was localized in a wrong room or on the wrong side of a partition, bookcase, or table within the same room. Figure 4.10 shows the cumulative distribution function of the localization error for both systems in the *Conference Room*—the area in the office environment where the greatest disparity in performance between the two systems was observed. Because there was only free space between sites on opposite sides of the table, the RF signatures there were too

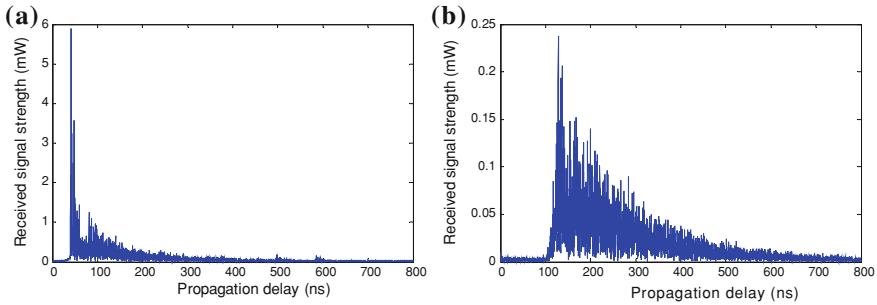


**Fig. 4.10** The cumulative distribution function of the localization error measured during testing in the *Conference Room*. Logical errors, denoted as X, were identified when the mobile was localized in a wrong room or on the wrong side of a partition, bookcase, or table within the same room

similar for robust discrimination between them. In fact, in the case where a memoryless system model is used ( $n = 1$ ), almost 50 % of the time it reported a logical error whereas in the case where a five-state sequence was used, the error was reported only 10 % of the time. Clearly, information about a single state alone does not suffice to correctly identify the trajectory; rather, information provided from multiple states taken collectively must be considered. This experiment underscores the strength of the Markov localization technique using sequences.

#### 4.4 Channel Impulse Response Fingerprinting

Thus far we have considered only the received signal strength feature for fingerprinting. RSS is typically used for localization in WLAN and 3G cellular networks (cellular localization systems are presented in [Chap. 5](#)). In these networks, signal strength is measured as the carrier power sensed over a period of time. While it varies by technology, the period is normally the duration of a packet. The power sensed is that of the transmitted signal arriving along the direct path from the base station. Also sensed are copies of the transmitted signal which arrive along other propagation paths. As explained in detail in [Chap. 2](#), the collection of copies is referred to as multipath. The multipath copies arrive delayed with respect to the direct path due to the characteristics of the propagation environment. When indexed according to delay, the power is referred to as the power delay profile or



**Fig. 4.11** **a** Channel impulse responses for a 6 GHz signal a line-of-sight (*LOS*) conditions. **b** non-line-of-sight (*NLOS*) conditions

the channel impulse response (CIR). Figure 4.11a illustrates the CIR for a 6 GHz signal and Fig. 4.11b illustrates another channel CIR from the same base station, however with the mobile displaced to a different location (more examples can be found in Chaps. 2 and 3). Notice the distinct properties between the two profiles. The profiles provide unique characterization of the mobile sites whereas the RSS signatures at the two sites—which is essentially just the integration of the power across the profile—may be very similar. As such, the channel impulse response can serve as an alternative fingerprint with enhanced identifiability. The CIR-fingerprinting technique was first introduced by US Wireless Corporation of San Ramon, California (Koshima and Hoshen 2000). Fingerprinting using the channel impulse response was also proposed for cellular UMTS localization (Ahonen and Eskelinen 2003a, b).

Channel impulse responses first appeared in localization in time-of-arrival based systems. Ideally, the first multipath arrival in the power delay profile will correspond to the direct propagation path between the base and the mobile. This then begs the question: if the channel impulse response is available, why not just use it to extract time-of-arrival? While TOA systems are capable of delivering accuracy on the order of several centimeters in line-of-sight conditions, the accuracy can degrade significantly in non-line-of-sight depending on the number, size, and material type of the obstacles between the radios. For example, in industrial environments rich in metal scatterers, deflection of the direct path off the straight line between the TX to the RX can cause a significant delay in the first arrival; or in subterranean mines, being that the walls are impenetrable by the direct path, the first arrival detected must necessarily correspond to some other path with a longer delay. In these cases, CIR fingerprinting offers a viable solution. In fact, radio frequency fingerprinting systems thrive on environments rich in scattering, such as the industrial environment. The scattering helps create distinctive signatures even between trained sites in close proximity. Since the multipath signature is unique and varies from one location to the next, it is even possible to implement a fingerprinting system with a single base station.

One benefit of survey-based techniques is that they can be implemented with existing infrastructure, necessitating no proprietary location and tracking equipment. For example, RSS fingerprinting systems exploit Wi-Fi base stations, which are both cheap and evermore ubiquitous worldwide, reducing deployment costs significantly. In the past, measuring the channel impulse response required expensive laboratory equipment such as the Vector Network Analyzer system described in [Chap. 2](#). Nowadays, channel impulse responses can be measured with complex receivers used in high-speed, wide-bandwidth systems. For instance, in wideband 4G cellular networks which use Orthogonal Frequency Division Multiplexing (OFDM), the frequency response of the channel, otherwise known as the Channel Transfer Function (CTF), is measured from the preamble of a packet in order to enable channel estimation and equalization. The channel impulse response can then be recovered from the CTF by converting it to the delay domain through the inverse Fourier Transform. The wider the bandwidth of the telecommunications system, the better its capacity to resolve multipath ([Gentile and Kik 2007](#)). In fact, because narrowband systems have poor resolution, the different arrivals appear as if they were grouped all as one. As a result, the power from the different paths cannot be discriminated and it is detected, rather, as a single quantity over the period, i.e. as the RSS value.

Two CIR-based systems are considered in this section. In [Sect. 4.4.1](#), a system which was implemented inside a mine tunnel is described. The system only processes the magnitude information of the multipath delay components. An improvement to the implementation, in which a nonparametric regression technique also exploits the phase information, is described in [Sect. 4.4.2](#).

### ***4.4.1 Mapping Using a Neural Network***

Since mine shafts are typically void of objects, they tend to have poor scattering properties. Then for the reasons explained earlier, received signal strength fingerprinting may deliver unacceptable resolution. As demonstrated in [Nerguizian et al. \(2006\)](#), channel impulse response fingerprinting in the mine environment, instead, can achieve good performance. In this subsection, we provide an overview of this paper. In the fingerprinting stage, the CIR was recorded at a number of sites throughout the mine using a Vector Network Analyzer, which acted as the sole base station<sup>3</sup>. In reference to [Eq. \(3.38\)](#), the channel impulse response can be represented mathematically as a train of uniformly sampled complex amplitudes,  $h(\tau_k)$ , indexed according to delay  $\tau_k$ ,  $1 \leq k \leq L_p$ . Each fingerprinted site was characterized by seven representative features extracted from the channel impulse response. The main features, which have already been introduced in [Chap. 3](#), are the received signal power

---

<sup>3</sup> Details of the VNA are described in [Chap. 2](#).

$$P = \sum_{k=1}^{Lp} |h(\tau_k)|^2, \quad (4.35)$$

the mean excess delay

$$\bar{\tau} = \frac{1}{P} \sum_{k=1}^{Lp} \tau_k \cdot |h(\tau_k)|^2, \quad (4.36)$$

and the root mean square delay spread

$$\tau_{\text{RMS}} = \sqrt{\frac{1}{P} \sum_{k=1}^{Lp} (\tau_k - \bar{\tau})^2 \cdot |h(\tau_k)|^2}. \quad (4.37)$$

The other features are the number of components  $L_{\text{SNR}}$  whose power is above a designated signal-to-noise ratio threshold  $T_{\text{SNR}}$ , i.e.  $L_{\text{SNR}} = \sum_{|h(\tau_k)|^2 \geq T_{\text{SNR}}} k$ , the time-of-arrival  $\tau_1 = \min_{|h(\tau_k)|^2 \geq T_{\text{SNR}}} \tau_k$  the power of the first arrival  $P_1 = |h(\tau_1)|^2$ , and the maximum arrival time  $\tau_{\text{MAX}} = \max_{|h(\tau_k)|^2 \geq T_{\text{SNR}}} \tau_k$ .

For the experiment in Nerguizian et al. (2006), close to 400 sites were fingerprinted throughout a surveyed mine. The multilayer perceptron, which was described in Sect. 4.2.3, was utilized to perform the mapping from the CIR-feature space to the location space. The seven features extracted for each of the sites, coupled with the two-dimensional location coordinates of each site, were used to train the perceptron. Accordingly, the neural network had seven inputs and two outputs. In this application, only one hidden layer with ten neurons was sufficient for training. The results for this method are presented in a side-to-side comparison in the next subsection.

#### 4.4.2 Mapping Using a Gaussian Kernel

In the work presented above, the seven features of the channel impulse response described were deemed sufficient to discriminate the sites throughout the survey area. Aside from the benefits of a compact representation, the authors in Jin et al. (2010) argue that, by extracting these features only, useful information available for location identification is discarded. In their paper, the authors show that by exploiting the unreduced CIR, results can be improved significantly. To this end, let the channel impulse response at site  $i$  from base station  $j$ , denoted as  $\mathbf{h}_{ij}$ , be a vector of  $L$  received power values sampled at uniform delay intervals. The signature at site  $i$  is then given through the collection of the CIR vectors from each of the  $n_B$  base stations; together they form the concatenated *supervector*  $\mathbf{H}_i = [\mathbf{h}_{i1} \mathbf{h}_{i2} \dots \mathbf{h}_{i,n_B}]^T$  of length  $n_B \times L$ .

Once the sites have been fingerprinted, the mobile location is estimated as a linear combination of the locations of the  $n_M$  trained sites:

$$\tilde{\mathbf{x}} = \frac{1}{n_M} \sum_{i=1}^{n_M} \rho_i \mathbf{x}_i \quad (4.38)$$

The weight associated with each site is the similarity metric—between the CIR supervector,  $\hat{\mathbf{H}}$ , measured during localization and the fingerprinted CIR supervector,  $\mathbf{H}_i$ . In the paper, the similarity metric is selected as the Gaussian kernel function

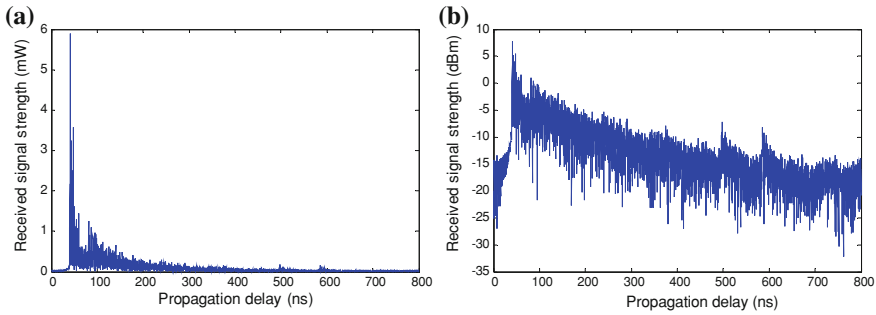
$$\rho_i = \frac{1}{\sqrt{(2\pi)^{n_M} |\Sigma|}} e^{-\frac{1}{2}(\hat{\mathbf{H}} - \mathbf{H}_i)^T \Sigma^{-1} (\hat{\mathbf{H}} - \mathbf{H}_i)}, \quad (4.39)$$

where  $\Sigma$  is the sample covariance matrix of the  $n_M$  fingerprinted supervectors. The sample covariance matrix is defined as

$$\Sigma = \frac{1}{n_M} \sum_{i=1}^{n_M} \mathbf{H}_i \cdot \mathbf{H}_i^T \quad (4.40)$$

In practice, fingerprinted sites throughout a survey area will be statistically correlated. For instance, adjacent sites collocated in a hallway will receive correlated multipaths—both in strength and in delay—from the same base station, especially those multipaths reflected into the hallway from the same direction. The key strength of the Gaussian kernel function as a similarity metric is that through the covariance matrix, the statistical correlation between the CIR supervectors of the fingerprinted sites is captured. This is a departure from the independence assumptions of Eq. (4.8) in Sect. 4.1.2.1 and of Eq. (4.24) in Sect. 4.2.4. Note that in Nerguizian et al. (2006), this statistical correlation is also captured, however, more implicitly through a neural network. Whereas in the latter the CIRs are processed on a linear scale, in Jin et al. (2010) the CIRs are processed on a logarithmic scale for the following reason. Since arrivals with larger delay are significantly attenuated with respect to the direct path, the logarithmic scale serves to leverage the contribution of each arrival; otherwise the contribution of the later arrivals would be dwarfed by the earlier arrivals. Figure 4.12 shows the same channel impulse response on a linear scale in (a) and on a logarithmic scale in (b).

We now present the results from the performance comparison in Jin et al. (2010). The paper compares the method described in this subsection, referred to as *LOG-ACIR-NKR*, to the method described in the previous subsection, referred to as *ACIR-GRNN*. Also compared was the *RSS-Kernel* method which draws on the same Gaussian kernel function in (4.39), however by replacing the CIR-supervector features,  $\hat{\mathbf{H}}$  and  $\mathbf{H}_i$ , with the RSS-vector features,  $\hat{\mathbf{P}}$  and  $\mathbf{P}_i$  (the RSS value,  $P$ , was computed from the generated CIR as in (4.35)). The three methods were implemented using raytracing software: given the three-dimensional CAD model of a building together with the thickness and the dielectric properties of the walls,



**Fig. 4.12** The channel impulse responses for a 6 GHz signal in line-of-sight conditions. **a** The signal power on a linear scale. **b** The signal power on a logarithmic scale

the software can generate high-resolution CIRs based on the configured positions of the base stations and the mobile device throughout the environment. The bandwidth of the system was set to 60 MHz and there were two base stations and 173 fingerprinted sites spaced at 1.5 m in the deployment area. Based on the cumulative distribution function of the location error for all three methods, the *LOG-ACIR-NKR* method achieves an average location error as low as 1 m, followed by the *ACIR-GRNN* method with an average error of 1.7 m, and then by the *RSS-Kernel* method with an error of 3.2 m. The *LOG-ACIR-NKR* method outperforms the *ACIR-GRNN* method mainly because it operates on a logarithmic scale as opposed to on a linear scale, making full use of distinctive properties of even the weakest arrivals in the CIRs. As expected, the *LOG-ACIR-NKR* method easily outperforms the *RSS-Kernel* method because it exploits the supplemental information provided by the CIR, with the arrivals sorted by delay rather than grouped as a single value.

### 4.4.3 Variations of CIR Fingerprinting

Analogous to the time domain channel impulse response, the channel transfer function can be used alternatively for fingerprinting. The CTF contains the same multipath channel information, however in the form of complex samples in the frequency domain. Similarly, the CTF correlation function, known as the Frequency Channel Coherence Function (FCF), is also proposed as a signature in (Malik and Allen Nov. 2006). The paper shows that the FCF is more stable and has superior performance to the CTF. A patent application proposes a similar technique that integrates FCF-based fingerprinting in existing OFDM-based systems (such as WLANs) (Bevan et al. 2010). Finally, the multipath characteristics alone can be further enhanced by incorporating an antenna array at the receiver. The antenna array enables the spatial characteristics, not just the temporal, of multipath—indexed according to both arrival angle and delay—to be captured. This results in a

richer signature defined as the power spatial delay profile (PSDP) (Triki et al. 2006; Gentile and Braga 2008).

## 4.5 Non-Radio Frequency Features

Thus so far we have considered only radio frequency features for survey-based location systems. Depending on which features are selected for a particular application—as well as other design parameters such as the number of base stations and the number of fingerprinted sites in a survey area—systems typically deliver localization accuracy between 1 and 3 m. For many applications, the expenses associated with the equipment and infrastructure necessary to deliver this level of accuracy are too high. For such applications, precise physical location within a certain environment is not required; rather, determining whether a mobile user lies within a confined space with a high degree of reliability takes priority. This type of *logical* localization is important in environments such as stores, museums, libraries, gas stations, etc. For instance, in a museum the appropriate automated tour can be offered as the visitor approaches the entry of an exhibition. In a grocery store, location-based services can notify a shopper of available coupons for items while walking along the aisles where they are stocked; in this environment, even if a device can deliver accuracy up to 2 m, this accuracy may not suffice to determine whether the mobile is in one aisle or the one adjacent to it.

Other features can be used to supplement, or even substitute, radio frequency fingerprinting. In this section, we investigate the application of the features described in Azizyan et al. (2009), namely the non-RF features of sound, motion, and color and the RF-feature of connectivity. For example, in a Laundromat the authors observed that sound is characterized by moving mechanical parts while in a library, on that other hand, it is very quiet. Analogously, typical motion in a supermarket involves walking up and down aisles with periodic pauses to select items; this contrasts static motion in a restaurant where the customer remains mostly seated for the duration of the stay. The chromatic features take advantage of the fact that many stores have trademark colors which are accentuated throughout the environment, such as red and white in *Target* © or pink and orange at *Dunkin' Donuts* ©. Finally, regarding connectivity, a mobile device can form a radio link only with base stations within the vicinity of the environment; hence, connectivity alone—as opposed to the degree of connectivity expressed by received signal strength—can be exploited as a signature for the environment. While the individual features may be similar from environment-to-environment, combining all four of the features together can prove to be highly discriminatory. The authors show that it is possible to achieve logical-localization accuracy of up to 87 % in 51 different environments using only these features which are accessible on most smartphones. In the remainder of this section, we described these four features in greater detail.



### 4.5.1 *Sound Features*

Sound in an environment is characterized through the temporal distribution of its volume intensity. This distribution is represented by a histogram of the intensity values recorded over a 1 min segment. The histogram is divided into 100 bins of equal size ranging from the minimum to maximum volume of the mobile device. In the Laundromat environment, for example, the histogram is very sharp at the center, indicating a constant buzz of medium intensity; this can be attributed to the rotation of the internal parts of the washers and dryers. Conversely, the distribution in a coffee shop is wider by virtue of the traits of human conversation, composed from a greater range of volumes—from the baristas preparing the items and calling out orders to conversational chatter in the background. The histogram vector serves as the signature for an environment. When compared against a vector measured during the localization stage, the inverse of the Euclidean distance—the distance is computed in the 100-dimensional space of the histogram vector—is used as the similarity metric. Notice that, as opposed to the previous sections in this chapter, the authors' convention in Azizyan et al. (2009) is that a larger similarity metric is more favorable.

### 4.5.2 *Motion Features*

Most smartphones now offer location services. When GPS is available, mainly outdoors and in some indoor environments—especially indoor environments with many windows through which the signal can penetrate—GPS can furnish the location of the mobile device. In GPS-denied areas, however, accelerometers can be employed to interpolate between the GPS readings (details of inertial-based systems are provided in Chap. 8). It turns out that in application to fingerprinting, accelerometers can be exploited to a second end in order to classify the types of motion which are common in an environment. This is accomplished by extracting signatures from the accelerometer readings. In Azizyan et al. (2009), each reading is the output of one of three-dimensional accelerometer axes. When sampled over time, two sequences are generated from each one of the three readings: one is the moving averages of the readings and one is their moving variances. The sequences are then fed to a Support Vector Machine (see Sect. 4.2.2) which classifies the sequences into one of two categories: either moving or stationary. The SVM is trained from the samples of the two sequences.

The actual signature is then computed as the quantity  $r$ , which is the ratio of time the mobile device is moving to the time it is stationary—over some observation window. The signature is then categorized into three classes: sitting for  $0 \leq r \leq 0.2$ , browsing for  $0.2 < r \leq 2.0$ , and walking for  $r > 2.0$ . By associating one or more classes to each of the fingerprinted environments, the signature can be used for the purpose of discriminating between the environments in which the

different classes occur. Then, the similarity metric between the signature measured during the localization stage and the signature of a fingerprinted environment is a value between 0 and 3. The value indicates the number of classes which the two signatures have in common.

### 4.5.3 Color Features

In order to capture the color of an environment, the floor is chosen as the target area. The reason is twofold: first of all the view of the floor (tiles, carpeting, wood, etc.) is relatively static over time, changing mainly due to obstructions from pedestrian walking. The ceiling would make for an even better candidate since it lacks such obstructions, however smartphones usually have their camera installed on the back of the device and so the camera is seldom pointed toward the ceiling. This makes for the second reason why the floor is chosen. Based on how the smart phone is held, the phone can discern one of its six possible orientations. From its orientation, the phone can determine at which times it is pointed towards the floor.

The camera's charge-couple device digitizes the captured image into RGB (Red–Green–Blue) pixels. Because the RGB chromatic space was found to be too sensitive to shadows and reflections, the pixels were transformed into the more robust HSL (Hue–Saturation–Lightness) space. The fingerprinting procedure consisted of the following steps. First the HSL pixels from all the images taken in the same environment were captured and grouped into clusters in the three-dimensional space via the *k-Nearest Neighbor* method (see Sect. 4.2.1). The number of clusters in any environment ranged from 3 to 7. Each cluster was represented by its HSL centroid and the signature of the environment was designated as the centroids of the clusters identified. Finally the similarity metric was computed as the distance between the centroids of the image(s) captured during localization and the centroids of environment  $i$ . Specifically, the similarity metric is expressed as:

$$\rho_i = \sum_k \sum_l \frac{1}{d_i(k, l)} \left( \frac{\hat{N}^k}{\hat{N}} \right) \left( \frac{N_i^l}{N_i} \right) \quad (4.41)$$

where  $d_i(k, l)$  is the Euclidean distance in the HSL space between centroid  $k$  of the captured image(s) and centroid  $l$  of environment  $i$ . The value  $\hat{N}^k$  indicates the number of pixels in cluster  $k$  and the value  $N_i^l$  indicates the number in cluster  $l$ . Analogously, the value  $\hat{N}$  indicates the total number of pixels over all the clusters of the captured image(s) while  $N_i$  indicates the total number over all the clusters of environment  $i$ . Note that each term in Eq. 4.41 corresponds to a cluster pair in the measured and fingerprinted spaces. Since the similarity metric is inversely proportional to the distance between the pair, the distance between the closest pair will have the greatest influence on the metric; likewise, the cluster pairs farthest apart will have the least influence. Each term is also weighted by the relative

number of pixels in the respective clusters of the pair. This downplays the contribution of smaller clusters which may just be outliers representing background noise.

#### 4.5.4 Connectivity Features

As the mobile device moves from one environment to another, its connectivity will change. Rather than measure the degree of connectivity between a base station and a mobile device through a similarity metric between their received signal strengths, a hard limit can also be used—that is—whether a radio link between the two can be established or not. Since the mobile device pings the surrounding base stations periodically in order to register their MAC addresses, the frequency of acknowledgment can be used as a similarity metric of connectivity instead. The authors quantify the frequency of the connectivity,  $f$ , as the fraction of acknowledgments the mobile receives from a particular base station to the total number it receives from all the  $n_B$  base stations during a fixed time period. The vector of connectivity values  $\mathbf{f}_i = [f_{ij}], j = 1 \dots n_B$ , that a mobile registers within environment  $i$  is designated as the signature of the environment. Then, the similarity metric between signature  $\hat{\mathbf{f}} = [\hat{f}_j], j = 1 \dots n_B$  measured during localization and signature  $i$  is:

$$\rho_i = \sum_{j=1}^{n_B} (\hat{f}_j + f_{ij}) \cdot \frac{\min(\hat{f}_j, f_{ij})}{\max(\hat{f}_j, f_{ij})} \quad (4.42)$$

Term  $j$  is large when the measured and fingerprinted connectivities to station  $j$  have a comparably high frequency; if they are not comparable then the min over max factor will be very small and will attenuate the weight of the term in the sum.

There are many ways in which the sound, motion, color, and connectivity features described in this section can be leveraged in order to estimate the mobile's location. For example, their individual similarity metrics can be weighted in a linear combination to formulate a comprehensive similarity metric<sup>4</sup>. Alternatively, the authors in Azizyan et al. (2009) chose to apply the features in the following manner. First of all, the sound, motion, and connectivity features were used sequentially to filter out unlikely environments. For each, an appropriate threshold value was set and any candidate environment with a corresponding similarity metric below that value was discarded. Once the initial filtering was performed, the color feature was selected to ultimately determine the location of the mobile as the one with the largest similarity metric among the remaining environments.

---

<sup>4</sup> This requires normalization of the individual similarity metrics such that each of their minimum and maximum values falls between 0 and 1, respectively.

## 4.6 Remarks

Some of the most practical fingerprinting techniques and system applications have been described in this chapter. While survey-based techniques may benefit from exploiting existing wireless infrastructure or from the deployment of low-cost nonproprietary equipment, they still suffer from the drawbacks of a required fingerprinting stage. One drawback is that the system cannot be deployed ad hoc, rather necessitating hours, weeks, or even months of training for large-scale networks, e.g., cellular. (Apropos, the following chapter investigates geolocation systems in cellular networks). Another drawback is that the radio frequency characteristics of the sites vary with any environmental changes. Such changes may arise from the movement of furniture, partitions, and any other objects, altering the path loss exponent and shadow fading in the environment; also, the addition or removal of base stations modifies the structure of the database. It is worth mentioning that some recent effort has been dedicated to research in automatic fingerprint training (Kim et al. 2010; Eleryan et al. 2011).

The drawbacks of fingerprinting preclude mission-critical applications for which localization services are vital, such as in emergency response and for firefighting in particular. Even if, in theory, a fingerprinting system could be trained in advance and its database updated automatically, the signature characteristics of the environment would change drastically during the rapid progression of a fire. Walls and floors may collapse and the water from the fire extinguishing hoses (which has the property of high RF reflectivity) and the ambient smoke are suspected to change the propagation characteristics of the environment. Other factors may be the flame itself as well as any dust produced from the deteriorating structure. Studies of these factors are currently underway at the National Institute of Standards and Technology.

## References

- A. Agiwal, P. Khandpar, H. Saran, LOCATOR: Location Estimation Systems for Wireless LANs. International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots, ACM, pp. 102–109, 2004
- S. Ahonen, P. Ekelinen, Performance Estimations of Mobile Terminal Location with Database Correlation in UMTS Networks. International Conference on 3G Mobile Communications Technologies, pp. 25–27, (2003a).
- S. Ahonen, P. Eskelinen, Mobile terminal location for UMTS. *Aerosp. Electron. Sys. Mag.* **18**(2), 23–27 (2003b).
- M. Azizyan, I. Constandache, R.R.Choudhury, Surroundsense: Mobile Phone Localization via Ambience Fingerprinting. International Symposium on Mobile Ad hoc Computing and Networking, ACM, pp. 261–272, 2009
- P. Bahl, V. Padmanadhan, RADAR: an in-building rf-based user location and tracking system. *IEEE INFOCOM* **2**, 775–784 (2000)

- R. Battiti, T.L. Nihat, A. Villani, Location-Aware Computing: A Neural Network Model for Determining Location in Wireless LANs. Department of Information and Communications Technology, Technical Report DIT-020083, University of Trento, Italy, 2002
- D.D. Bevan, L. Averin, D. Lysyakov, RF Fingerprinting for Location Estimation 0311436, United States, 2010
- N.K. Bose, *Neural Network Fundamentals with Graphs, Algorithms, and Applications* (McGraw-Hill, New York, 1995)
- M. Brunato, R. Battiti, Statistical learning theory for location fingerprinting in wireless LANs. *Comp. Netw. ISDN Syst.* **47**(6), 825–845 (2005). Elsevier
- W. Burgard, et al., Estimating the Absolute Position of a Mobile Robot Using Position Probability Grids Conference on Artificial Intelligence, AAAI, pp. 896–901, 1996
- G. Cybenko, Approximations by superpositions of sigmoid functions. *Math. Control, Signals, Sys.* **3**(4), 303–314 (1989). Springer
- A.M. Edgar, C. Raul, F. Jesus, Estimating user location in a WLAN using back propagation neural networks. *Lect. Notes Comp. Sci.* **3315**, 737–746 (2004)
- A. Eleryan, M. Elsabagh, M. Youssef, AROMA: Automatic Generation of Radio Maps for Localization Systems. International Conference on Mobile Computing and Networking, ACM, pp. 93–94, 2011
- S-H. Fang, T-N. Lin, K-C. Lee, A novel algorithm for multipath fingerprinting in indoor WLAN environments, *Trans. Wireless Commun. IEEE*, 7(9) 2008
- D. Fox, W. Burgard, S. Thrun, Markov localization for mobile robots in dynamic environments. *J. Artif. Intell.* **11**, 391–427 (1999)
- D. Fox et al., Bayesian filtering for location estimation. *Pervasive Comput. IEEE.* **2**(3), 24–33 (2003)
- C. Gentile, A.J. Braga, A comprehensive evaluation of joint range and angle estimation in indoor ultrawideband location systems. *EURASIP J. Wireless Commun. Networking* **2008**, 248509 (2008). Hindawi
- C. Gentile, A. Kik, A comprehensive evaluation of indoor ranging using ultra-wideband technology. *EURASIP J. Wireless Commun. Networking* **2007**, 86031 (2007). Hindawi
- C. Gentile, L. Klein-Berndt, Robust Location Using System Dynamics And Motion Constraints. International Conference on Communications, IEEE. Paris, France, pp. 1360–1364, 2004
- C. Gentile, S.M. Lopez, A.A. Kik, Comprehensive spatial-temporal channel propagation model for the ultra-wideband spectrum 2–8 GHz. *IEEE Trans. Antennas Propag.* **58**(6), 2069–2077 (2008)
- V. Honkavirta et al., A comparative survey of wlan location fingerprinting methods. Workshop on Positioning, Navigation and Communication. pp. 243–251, March 2009
- V. Honkavirta, Location fingerprinting methods in wireless local area networks. Master of Science Thesis, Tampere University of Technology, Finland, Oct 2008
- M. Isard, A. Blake, condensation: conditional density propagation for visual tracking. *Int. J. Comput. Vision* **1**, 5–28 (1998)
- Y. Jin, W.-S. Soh, W.-C. Wong, Indoor localization with channel impulse response based fingerprint and nonparametric regression. *IEEE Trans. Wireless Commun.* **9**(3), 1120–1127 (2010)
- K. Kaemarungsi, P. Krishnamurthy, Modeling of Indoor Positioning Systems based on Location Fingerprinting. *INFOCOM. IEEE*, pp. 1012–1022, 2004
- R.E. Kalman, A new approach to linear filtering and prediction problems. *Trans. AMSE: J. Basic Eng.* **82**, 35–45 (1960)
- Y. Kim, Y. Chon, H. Cha, Smartphone-based collaborative and autonomous radio fingerprinting. *IEEE Trans. Syst. Man, Cybern. Part C: Appl. Rev. IEEE*, **2**(1), 112–122 (2010)
- H. Koshima, J. Hoshen, Personal locator services energy. *Spectrum.* **37**(2), 41–48 (2000)
- A. Kushki, K.N. Plataniotis, A.N. Venetsanopoulos, Kernel-based positioning in wireless local area networks. *Mobile Computing.* **6**(6), 689–705 (2007)
- A.M. Ladd et al., Robotics-based location sensing using wireless ethernet. *Wirel. Networks* **11**, 189–204 (2005). (Springer Science + Business Media, Inc.)

- Z. Li Wu et al., Location estimation via support vector regression. *Trans. Mobile Comput. IEEE* **6**(3), 311–321 (2007)
- T.-N. Lin, P.-C. Lin, Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks. *Wireless Netw., Commun.Mobile Comput.* **2**, 1569–1574 (2005)
- H. Liu et al., Survey of wireless indoor positioning techniques and systems. *IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev.* **37**(6), 1067–1080 (2007)
- D. Madigan et al., Bayesian Indoor Positioning Systems, vol. 2. *INFOCOM. IEEE*, pp. 1217–1227, 2005
- W.Q. Malik, B. Allen, Wireless Sensor Positioning with UWB Fingerprinting. *European Conference on Antennas and Propagation*, pp. 1–5, Nov 2006
- Y. Moustafa, A. Ashok, The Horus WLAN Location Determination System. *International Conference on Mobile Systems, Applications and Services*, pp. 205–218, 2005
- C. Nerguizian, C. Despins, S. Affes, Geolocation in mines with an impulse response fingerprinting technique and neural networks. *IEEE Trans. Wireless Commun.* **5**, 603–611 (2006)
- T. Roos et al., A probabilistic approach to WLAN user location estimation. *Int. J. Wireless Inf. Networks* **9**, 155–164 (2002)
- A.J. Smola, B. Schoelkopf, A tutorial in support vector regression. *Stat. Comput.* **14**, 199–222 (2004). (Kluwer Academic Publishing)
- M. Triki et al., Mobile Terminal Position via Power Delay Profile Fingerprinting: Reproducible Validation Simulations. *Vehicular Technology Conference, IEEE*, pp. 1–5, Sept 2006
- C.L. Wu, L.C. Fu, F.L. Lian, WLAN Location Determination in e-Home via Support Vector Classification. *International Conference on Networking, Sensing and Control, IEEE*, pp. 1026–1031, 2004
- M.A. Youssef, A. Agrawala, A.U. Shankar, WLAN Location Determination via Clustering and Probability Distribution. *International Conference in Pervasive Computing and Communications, IEEE*, pp. 143–150, 2003