CHANNEL LEARNING IN INDOOR LOCALIZATION

Mohamed F. Mansour

Texas Instruments Inc.
Dallas, Texas, USA

ABSTRACT

This work addresses the fundamental issue of online channel learning in indoor localization systems. We describe a novel algorithm for unsupervised channel learning where the path loss exponent in the Non-Line-Of-Sight (NLOS) region is adapted online while positioning.The algorithm has scalable complexity and is shown to provide significant improvement over non-adaptive systems.

***Index Terms*—** Indoor localization; WLAN; MMSE; RSS; indoor channel model; EM algorithm; channel estimation.

1. INTRODUCTION

Most proposals of indoor localization systems use the received signal strength (RSS) measurements from multiple access points (AP) because of its relatively low-cost hardware requirement [1, 2, 3]. In WLAN-based positioning, the indoor channel is mostly NLOS except at distances close to the access points. A fundamental problem in these systems is characterizing the wireless channel between the reference points and the receiver so as to properly map the RSS to distance. Most indoor positioning systems use the log-normal shadowing model [4, 5] in which the RSS (in dBm) is proportional to the logarithm of the distance between the transmitter and the receiver. We adopt the more general two-region log-normal shadowing model [6], that describes a piecewise linear relation between the RSS (in dBm) and the distance. The model uses two different deterministic values for the path-loss exponent in the LOS and the NLOS regions. We further generalize the model by treating the NLOS path loss exponent as a random variable with a normal distribution whose mean is dependent on the indoor environment.

The main contribution of this work is developing an algorithm for *online* adaptation of the NLOS path loss exponent. It is an unsupervised procedure that uses the positioning estimate from the Minimum Mean-Square Error (MMSE) position estimator to adapt the channel model to the observed RSS measurements. The adaptation algorithm has an EM-like structure, where we alternate between estimating the position and estimating the mean of path loss exponent. It uses the positioning estimates over a large time window to reduce the impact of spurious positioning errors.

Many earlier RSS-based positioning systems, e.g., [7, 8], assumed perfect knowledge of the indoor channel. This assumption is justified if the system is used in a fixed and controlled environment, but cannot be extended to a general solution. In some other systems, e.g. [9], explicit estimation of the path loss exponent is avoided by searching over its admissible range for the best grid point. Explicit online estimation of the path loss exponent has been proposed in some earlier works, e.g., [10, 11, 12]. In some cases, e.g., [10, 11], it is assumed that all channels have the same path loss exponent over a period of time. This simplifies the problem to a joint estimation problem that is solved by nonlinear least square techniques. In [12], the path loss exponents are treated as unknown deterministic variables and are estimated from the RSS measurement using measures of compatibility of the distances estimates. This results in a nonlinear least square problem with the a search space whose size is proportional to the number of AP's. The proposed channel learning algorithm differs from prior art in the following aspects:

1. The path loss exponent in the NLOS region is treated as a random variable, and different channels can have different path loss exponents.
2. The procedure does not need offline calibration and the channel adaptation is performed online.
3. A generalized two-region channel model is used; which is more appropriate in characterizing the RSS in WLAN systems [6].
4. The localization problem is decoupled from the channel learning problem and both the observed and unobserved AP's are used for estimation as in [13].

Throughout the paper, we use the bold lower-case letters denote column vectors. x_k denotes the k-th element of $\mathbf{x}$.

2. CHANNEL MODEL

The indoor propagation channel is usually modeled using the log-normal distribution model [4, 5]. The received signal strength (in dBm) from the k-th acces point is modeled as

$$r_k(\text{dBm}) = \eta_k - 10\gamma_k \log_{10}\left(\frac{d_k}{d_0}\right) + w_k \qquad (1)$$

where w_k is a zero-mean Gaussian noise with variance σ^2, γ_k is the path-loss exponent, d_k is the Euclidean distance between the receiver and the k-th access point and η_k is the reference power (in dBm) at distance d_0. In practice, the value of γ_k varies with the indoor environment and the operating frequency band. Therefore, γ_k cannot in general be treated as a deterministic variable. Rather, it has a probability density function (PDF) that is determined by the indoor environment. This PDF could be approximated by a normal distribution with a mean that is dependent on the indoor environment. In Fig. 1, we show an example of the distribution of the path loss exponent in a typical indoor office environment that has both open areas with cubicles and high-wall rooms.

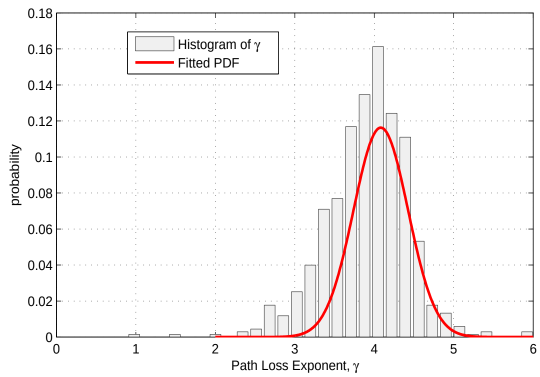

Fig. 1. An example of the distribution of the path loss exponent of NLOS channels in a typical indoor office environment.

We adopt the two region channel model [6], where small distances between the AP and the Receiver Unit (RU) are assumed to have an LOS channel, whereas large distances are assumed to have an NLOS channel. The border between small and large distances is a predetermined distance d_1 which is set in the range 5-10 meters from the AP. The channel model is characterized by the mean of the log-normal shadowing model as [6]

$$\mu_k = \begin{cases} \eta_k - 10\gamma_0 \log_{10}(d_k/d_0) & \text{if } d_k \le d_1 \\ \tilde{\eta}_k - 10\gamma_k \log_{10}(d_k/d_1) & \text{if } d_k > d_1 \end{cases} \tag{2}$$

where γ_0 is the LOS path loss exponent, and $\tilde{\eta}_k$ is chosen such that the continuity is preserved at $d = d_1$ (at any value of γ_k).

We assume that the RSS measurements from the different AP's are independent and have the same variance. The impact of fading can be mitigated by proper scanning strategies and averaging before processing the RSS measurements [14].

3. THE MMSE ESTIMATOR

It was shown in [13] that the incorporation of the unobserved AP's significantly improves the overall performance especially under poor AP geometry. An AP is not observed if the corresponding RSS is lower than the receiver sensitivity, λ. The log-likelihood function if both observed and unobserved access points are used in estimation is [13]

$$L(\mathbf{r}, \mathcal{A}|x, y, \boldsymbol{\gamma}) = \sum_{k\in\mathcal{A}} \frac{-(r_k - \mu_k)^2}{2\sigma^2} + \sum_{k\in\overline{\mathcal{A}}} \log Q\left(\frac{\mu_k - \lambda}{\sigma}\right) \tag{3}$$

where $\mathcal{A}$ and $\overline{\mathcal{A}}$ denote respectively the sets of indices of observed and unobserved access points; and $Q(t) \triangleq \int_t^\infty e^{-x^2/2}$ [15].

If $\{\gamma_k\}$ are unknown, we need to jointly estimate $(x, y; \{\gamma_k\})$ in (3). The straightforward approach for estimating (x, y) is to treat $\{\gamma_k\}$ as nuisance parameters and average it out through its PDF. Based on our earlier discussion, the distribution of γ_k in the NLOS region can be approximated as

$$\gamma_k \sim \mathcal{N}(\overline{\gamma}, \delta^2) \quad \text{for all } k \tag{4}$$

and $\{\gamma_k\}$ are independent. For consistency, we assume that γ_k in the LOS region has another normal distribution with $\overline{\gamma} = \gamma_0$, and $\delta = 0$. Defining the average likelihood as

$$\overline{L}(\mathbf{r}|x, y) \triangleq \int_{\boldsymbol{\gamma}} L(\mathbf{r}, \mathcal{A}|x, y, \boldsymbol{\gamma}) f_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}) \tag{5}$$

Then after substituting from (3) and discarding the irrelevant terms, we get

$$\overline{L}(\mathbf{r}|x, y) = \sum_{k\in\mathcal{A}} \frac{r_k}{\sigma^2} \int_t \mu_k f_{\gamma_k}(t) - \frac{1}{2\sigma^2} \int_t \mu_k^2 f_{\gamma_k}(t) + \sum_{k\in\overline{\mathcal{A}}} \int_t \log Q\left(\frac{\mu_k - \lambda}{\sigma}\right) f_{\gamma_k}(t) \tag{6}$$

To simplify notations, we define for the NLOS region

$$\begin{aligned} \overline{\mu}_k(x, y) &\triangleq \int_t \mu_k(x, y, t) f_{\gamma_k}(t) \\ &= \tilde{\eta}_k - 10\overline{\gamma} \log_{10}(d_k/d_0) \end{aligned} \tag{7}$$

$$\begin{aligned} \omega_k(x, y) &\triangleq \int_t \mu_k^2(x, y; t) f_{\gamma_k}(t) \\ &= \tilde{\eta}_k^2 - 20\tilde{\eta}_k\overline{\gamma} \log_{10}(d_k/d_0) + \\ &\qquad 100(\delta^2 + \overline{\gamma}^2) \log_{10}^2(d_k/d_0) \end{aligned} \tag{8}$$

Note that, $\log Q(x)$ can be approximated by a piecewise-linear function. Hence,

$$\int_t \log Q\left(\frac{\mu_k(x, y; t) - \lambda}{\sigma}\right) f_{\gamma_k}(t) \approx \log Q\left(\frac{\overline{\mu}_k(x, y) - \lambda}{\sigma}\right) \tag{9}$$

Hence the revised likelihood function in (6) becomes

$$\overline{L}(\mathbf{r}, \mathcal{A}|x, y) = \sum_{k\in\mathcal{A}} \frac{r_k \overline{\mu}_k(x, y)}{\sigma^2} - \frac{\omega_k(x, y)}{2\sigma^2} + \sum_{k\in\overline{\mathcal{A}}} \log Q\left(\frac{\overline{\mu}_k(x, y) - \lambda}{\sigma}\right) \tag{10}$$

This likelihood function is evaluated at a set of candidate points distributed over a grid whose center is the centroid of the observed access points. The maximum likelihood estimate is the grid point that corresponds to the maximum value.

The Minimum Mean-Square Error (MMSE) estimator can be evaluated similarly. If the unknown position (x, y) is treated as a random variable, the MMSE estimator is

$$\begin{aligned}\widehat{\mathbf{g}} &\triangleq \int_x \int_y (x, y)^T f(x, y|\mathbf{r}) \\ &= \frac{1}{f(\mathbf{r})} \int_x \int_y (x, y)^T f(\mathbf{r}|x, y) f_a(x, y) \end{aligned} \quad (11)$$

where $f_a(x, y)$ is the apriori PDF of the position and it can have several forms that are described in details in [16]. In the simplest case, it has a uniform distribution over the search grid. In this case, the MMSE estimator becomes

$$\widehat{\mathbf{g}} = c \sum_x \sum_y (x, y)^T \exp\left(\overline{L}(\mathbf{r}|x, y)\right) \quad (12)$$

where $\overline{L}(\mathbf{r}|x, y)$ is as in (10) and c is a normalization factor.

The inexact knowledge of all $\{\gamma_k\}$ causes performance degradation of RSS-based indoor localization systems that use an AP database when compared to fingerprinting-based systems (where a good approximation of each γ_k can be computed from the RSS fingerprint). This performance degradation is illustrated in Fig. 2 where we show the performance of the MMSE algorithm when all $\{\gamma_k\}$ are known (which approximates fingerprinting-based systems) versus the more general case when only $\overline{\gamma}$ is known. We also include the Modified Cramer-Rao Bound (MCRB) as described in [13]

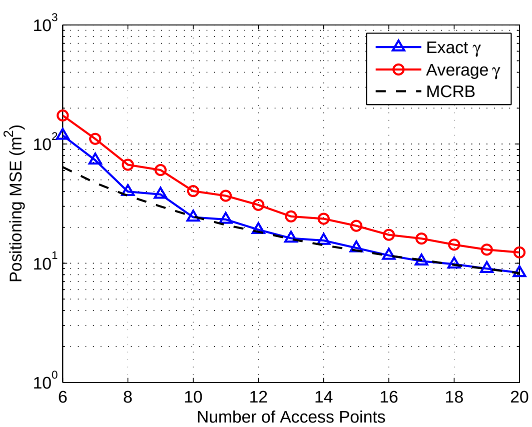

Fig. 2. Performance degradation of the MMSE algorithm from exact to average γ_k over an area of 100×100 m with $\overline{\gamma} = 4.1, \sigma = 3$ and $\delta = 0.3$

4. LEARNING ALGORITHM

The above analysis assumes that $\overline{\gamma}$ is deterministic and known. In practice, $\overline{\gamma}$ varies with the *terrain* of the indoor environment, and a versatile indoor positioning system must adapt to the indoor environment in order to operate reliably under all conditions. In Fig. 3 we show the impact on performance when incorrect $\overline{\gamma}$ is used, where different incorrect values are used rather than the true value. This highlights the need for the online learning of the channel during positioning.

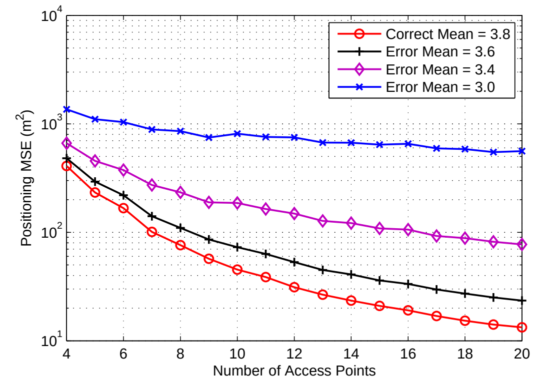

Fig. 3. Performance degradation of the MMSE algorithm with incorrect $\overline{\gamma}$ over an area of 100×100 m with $\overline{\gamma} = 3.8, \sigma = 3$ and $\delta = 0.3$

The objective of online channel learning is to estimate $\overline{\gamma}$ in the NLOS region from the RSS observations. If γ_k is considered as a random variable, then its MMSE estimator is,

$$\begin{aligned}\widehat{\gamma} &= E\{\gamma|\mathbf{r}\} \\ &= \int_x \int_y E\{\gamma|\mathbf{r}, x, y\} f(x, y) \end{aligned} \quad (13)$$

where $f(x, y)$ is the PDF of the receiver position. In the NLOS region, we have from (2)

$$E\{\gamma_k|r_k, x, y\} = \frac{\tilde{\eta}_k - r_k}{10 \log_{10}(d_k(x, y)/d_1)} \quad (14)$$

where $d_k(x, y)$ is the distance between the point (x, y) and the k-th AP. The corresponding estimation variance is

$$\xi_k^2 = \frac{\sigma^2}{(10 \log_{10}(d_k(x, y)/d_1))^2} \quad (15)$$

Therefore, if we have N RSS observations, we get the Wiener filter relation

$$E\{\gamma|\mathbf{r}, x, y\} = \frac{1}{\nu} \sum_{k=1}^{N} \frac{1}{\xi_k^2} E\{\gamma_k|r_k, x, y\} \quad (16)$$

$$= \frac{10}{\nu\sigma^2} \sum_{k=1}^{N} (\tilde{\eta}_k - r_k) \log_{10}\left(\frac{d_k(x, y)}{d_1}\right) \quad (17)$$

where $\nu \triangleq \sum_{k=1}^{N} \xi_k^{-2}$. The PDF $f(x, y)$ in (13) can be approximated by a probability distribution over the grid of search points that is proportional to the corresponding likelihood function from the MMSE position estimator in (10).

If $\overline{\gamma}$ is not time-varying, then the estimation variance could be reduced by temporal averaging; which reduces the impact of spurious observations. Let $\widehat{\gamma}_t$ denote the estimate at time t, then the channel estimation formula becomes

$$\widehat{\gamma}_{t+1} = (1 - \epsilon)\widehat{\gamma}_t + \epsilon E\{\gamma|\mathbf{r}_t\} \quad (18)$$

where ϵ is a smoothing factor (typically ≤ 0.1), and $\mathbf{r}_t$ is the vector of RSS observations at t.

5. SIMULATION RESULTS

We use the following experimental setup for evaluating the proposed algorithms. The access points are randomly scattered over an area of 100×100 meters (according to a uniform distribution). The RU moves within the area at a speed of 1 m/s and the scanning rate is 1 Hz. The channel model between the receiver and each of the access points is as described in (1) and (2) with $\sigma = 3$ dBm. The path loss exponent in the NLOS region is distributed as $\mathcal{N}(\overline{\gamma}, \delta^2)$, with $\delta = 0.3$ and $\overline{\gamma} = 4.1$. The MMSE as described in (12) is evaluated on a 2D grid within the area with spacing of 2 meters between adjacent grid points. For each experiment, we evaluate 10^3 routes along the area, where each route has on average 100 scans.

First, we illustrate the effectiveness of the channel learning algorithm. In Fig. 4, we show an example of a typical learning curve of $\overline{\gamma}$. The true value is $\overline{\gamma} = 4.1$ and the initial value is 3.4. In the learning algorithm, we set $\epsilon = 0.01$ in (18). The figure has two cases: the first case is a hypothetical case; which assumes that the correct position is always known, while the second case uses the MMSE position estimate to update $\overline{\gamma}$. As illustrated in Fig. 4, the practical case converges properly and the joint estimation of the position and $\overline{\gamma}$ is similar to the case with known position.

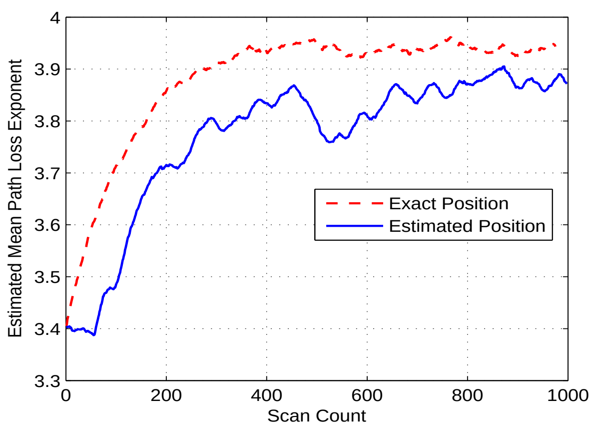

Fig. 4. Example of the learning curve for estimating $\overline{\gamma}$ with 20 access points

Next, we show the performance of the MMSE algorithm with channel learning for the above setup. The initial estimate of $\overline{\gamma}$ is set at 3.4 (recall the true value is 4.1). The performance is summarized in Fig. 5 where the MSE is averaged along the time of learning. As noticed from the figure, the performance improves significantly with learning. The standard deviation of the positioning error is around 10 meters with only 10 reference access points in a 100×100 meters area. Unlike most other positioning algorithms, the proposed algorithm adapts to the indoor environment in an unsupervised fashion and this performance is universal to any indoor environment that has the underlying general probabilistic model for the indoor channel.

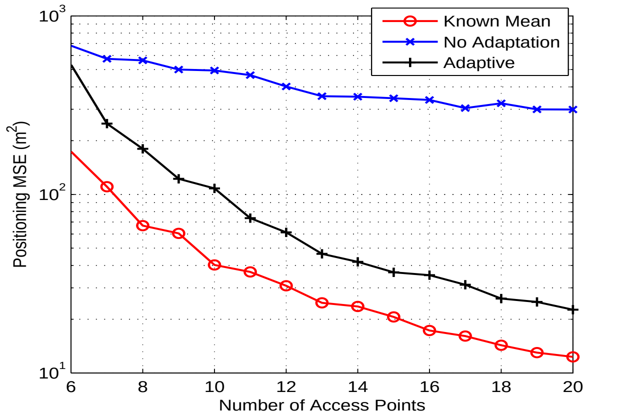

Fig. 5. Performance of the channel learning algorithm when $\overline{\gamma} = 4.1$ and the initial $\gamma = 3.4$

6. CONCLUSION

To enable *universal* indoor positioning, the system must adapt to the variability of the communication channel, which is inevitable in the indoor environment. The NLOS path loss exponent in the indoor channel could vary widely between 2-6, and it is also dependent on the operating frequency band. This work addresses this fundamental problem with the following steps:

1. A generalized channel model, that combines both LOS and NLOS channels, is used. It treats the NLOS path loss exponent as an unknown random variable. This model encompasses other typical indoor channel models as special cases.

2. For this generalized model, we derived a novel MMSE estimation procedure that uses both observed and unobserved access points in estimation. It has scalable complexity and it is suited for embedded systems.

3. We proposed an unsupervised channel learning procedure that operates online while positioning. We showed that it provides significant improvement over non-adaptive systems especially when a sufficient number of access points is available.

The missed piece in this universal positioning system, that was not discussed here, is the outlier rejection mechanism, which excludes the access points that does not comply with the underlying channel model. This mechanism is described in details in [17].

The proposed positioning algorithms are memoryless, and the MMSE algorithm searches over a set of grid points that are equally probable. Future work will study the use of a Kalman filter to improve both the MMSE performance and the adaptation algorithm.

7. REFERENCES

[1] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche, "A comparative survey of WLAN location fingerprinting methods," *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on*, pp. 243–251, March 2009.

[2] Hui Liu, H. Darabi, P. Banerjee, and Jing Liu, "Survey of wireless indoor positioning techniques and systems," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[3] F. Seco, AR. Jimenez, C. Prieto, J. Roa, and K. Koutsou, "A survey of mathematical methods for indoor localization," *Intelligent Signal Processing, 2009. WISP 2009. IEEE International Symposium on*, pp. 9–14, Aug 2009.

[4] T. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, second edition, 2002.

[5] H. Hashemi, "The indoor radio propagation channel," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 943–968, Jul 1993.

[6] "IEEE p802.11 wireless lan TGn channel models," *IEEE 802.11-04/940r2*, Jan 2004.

[7] G. Chandrasekaran, M. A. Ergin, J. Yang, S. Liu, Yingying Chen, M. Gruteser, and R. Martin, "Empirical evaluation of the limits on localization using signal strength," *Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON '09. 6th Annual IEEE Communications Society Conference on*, pp. 1–9, June 2009.

[8] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 8, pp. 2137–2148, Aug 2003.

[9] J. Shirahama and T. Ohtsuki, "RSS-based localization in environments with different path loss exponent for each link," *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pp. 1509–1513, May 2008.

[10] Xinrong Li, "RSS-based location estimation with unknown pathloss model," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 12, pp. 3626–3633, December 2006.

[11] H. Laitinen, S. Juurakko, T. Lahti, R. Korhonen, and J. Lahteenmaki, "Experimental evaluation of location methods based on signal-strength measurements," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 1, pp. 287–296, Jan 2007.

[12] S. Mazuelas, A. Bahillo, R.M. Lorenzo, P. Fernandez, F.A F.A. Lago, E. Garcia, J. Blas, and E.J. Abril, "Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 821–831, Oct 2009.

[13] M. Mansour and D. Waters, "Uncensored indoor positioning," *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 824–828, July 2014.

[14] D. Waters, M. Mansour, and A. Xhafa, "WLAN scanning strategies for RSSI-based positioning," *IEEE Global Telecommunication conf., GLOBECOM*, pp. 493–497, 2013.

[15] J. Proakis and M. Salehi, *Communication Systems Engineering*, Pearson Education, second edition, 2002.

[16] M. Mansour, "A priori information in indoor positioning," *Pending US Patent, USPTO*, 2013.

[17] M. Mansour, "Apparatus and method for selecting access points for use in positioning," *US Patent No. 8,478,298; USPTO*, 2013.