Handover Delay in Cellular Wireless Systems

Sennur Ulukus*
Wireless **I**nformation **N**etwork **Lab**oratory
Rutgers University
P.O. Box 909
Piscataway, NJ 08855
732-445-4447, 732-445-3693 (FAX)

Gregory P. Pollini
Bellcore
NVC 3X-361
331 Newman Springs Road
Red Bank, NJ 07701
732-758-5672, 732-758-4372 (FAX)

Abstract

Handover is the process by which a mobile terminal maintains a communication session as it crosses through the coverage area of the grid of base stations providing service in a cellular wireless system. Because of imperfect channel measurements, it is difficult to determine the correct point for a handover. Acting too soon may force a handover back to the original base station; delaying too long may make it impossible for the mobile to meet its quality of service objectives. It is important to quantify the handover delay in terms of various system parameters in order to efficiently design cellular wireless communication systems.

This paper presents an exact expression for handover delay as a function of the base station separation, the standard deviation of log-normal shadowing, and the averaging distance. It also presents a simple closed form approximation that includes the hysteresis margin. These are the first such expressions available within the literature. The approximation is validated via simulation and is shown to work well over the range of interest.

1 Introduction

In a wireless communication system, the strength of the signal received from a base station at a mobile terminal varies due to the change in the distance of the mobile from the base station and the random nature of the communication channel. When the communication channel between the mobile and its current base station becomes poor, the mobile switches to a stronger base station and this event is called *handover*. Figure 1 shows a mobile moving from one base station (current base) to another (target base). Several *handover criteria* have been proposed in the literature. These are summarized below; see also [4] for an extensive listing of previous work on handover.

Relative Signal Strength Based Handover: The mobile communicates with the base station with the strongest averaged received signal. In Figure 1, the handover occurs at position A. Although simplest and easiest to analyze, this method is shown to cause too many unnecessary handovers. [2]

*This work was completed while the author was a consultant in the Wireless Systems Research Department of Bellcore, Red Bank, NJ.

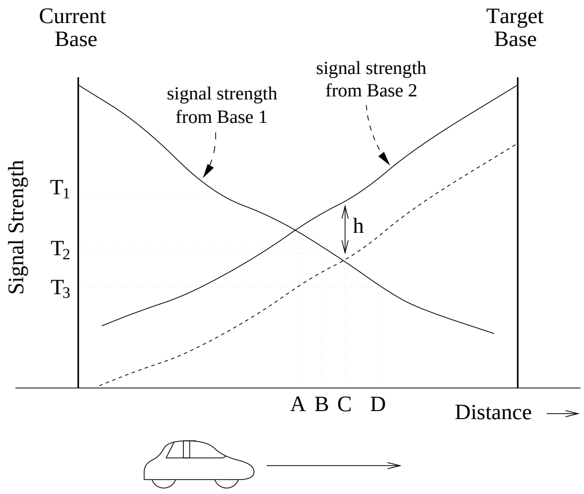

Figure 1: The location of a handover decision.

Relative Signal Strength with Threshold: The mobile chooses to handover to the new base station only if the current base station is sufficiently weak (the received signal is less than a fixed threshold) and the new one is the stronger of the two. The effect of the threshold depends on its value as compared to the signal strengths of the two base stations at the point at which they are equal. If the threshold is higher than this value, say T_1 in Figure 1, the scheme degenerates to the *Relative Signal Strength Based Scheme* and the handover occurs at A. If the threshold is lower than this value, say T_2 in Figure 1, the mobile delays handover until the current signal level crosses the threshold at position B. In the case of T_3, the delay may be so long that the mobile drifts far into the new cell. This reduces the quality of the communication link and may result in a dropped call. A threshold is not used as the lone metric in practice since its effectiveness depends on prior knowledge of the crossover signal strength between the current and candidate base stations.

Relative Signal Strength and Hysteresis Based Handover: The mobile changes its base station only if the received signal strength from the new base station is stronger by a margin (called a hysteresis margin) than the one from the current base station. In this case the handover occurs at

C. This technique prevents the so called ping-pong effect, the repeated handover between two base stations caused by rapid fluctuations in the received signals.

Relative Signal Strength with Hysteresis and Threshold: The mobile switches to the new base station only if the current base is weak enough and the candidate base is stronger than the current one by a hysteresis margin. In Figure 1, the handover occurs at point C if the threshold is either T_1 or T_2, and occurs at point D if the threshold is T_3.

The performance of a handover algorithm can be measured in terms of two metrics: expected number of handovers and handover delay, defined as the position at which the assignment probabilities of a mobile to both (current and target) base stations are equal. Since each handover event uses network resources such as the bandwidth required for signaling, it is desirable to design handover algorithms so that the expected number of handovers is as small as possible. Given the fixed system parameters: base station separation (D), standard deviation and correlation of log-normal shadowing (σ_s and d_o), path loss exponent ($K_2/10$), only three design parameters can be controlled: the averaging distance (d_{av}), the hysteresis level (h), and the threshold level (T). It is intuitively clear that increasing d_{av}, h and/or decreasing T decreases the expected number of handovers. Unfortunately, these actions increase the handover delay, resulting in a reduction in call quality and an increased probability of call dropping. Handover delay is a measure of the effective cell boundary; large handover delay is equivalent to dragging of the cell boundary towards the target base station and increases cell overlap areas.

Due to this trade-off between the expected number of handovers and the handover delay, it may be advantageous to select control variables (d_{av}, h and T) such that both performance metrics have acceptable values, given fixed system parameters. In order to quantify the trade-off between the expected number of handovers and the handover delay, closed form expressions are needed for these two quantities. In [5], a closed form approximation for the expected number of handovers is presented.

This paper develops a closed form approximation for the handover delay. The following section presents the propagation model and the necessary background on handover. Section 3 investigates the handover delay when the hysteresis margin is not used, $h = 0$. Section 4 adds the hysteresis margin. A summary is given in Section 5.

2 System Model

This paper considers a system where a mobile unit moves from one base station towards another on a straight line with a constant speed as shown in Figure 1. The received signal strengths from the two base stations are monitored by the mobile unit. The signals from the base stations are affected by two major factors: path loss and log-normal shadow fading. Rayleigh fading is neglected since we assume that the rapid fluctuations in the signals are averaged out. The received signal strength from base stations A (current) and B (target) are sampled at distances kd_s, k a positive integer, as a mobile moves away from base A. The received signal strengths in dB are denoted a_k and b_k, respectively, and are given as

$$\begin{aligned} a_k &= K_1 - K_2 \log(kd_s) + u(kd_s), \\ b_k &= K_1 - K_2 \log(D - kd_s) + v(kd_s), \end{aligned} \tag{1}$$

where u and v are zero mean Gaussian random variables that model the log-normal shadowing, D is the base station separation, K_1 is the signal strength at a distance $d = 1$, and K_2 is ten times the path loss exponent. Since threshold levels are not considered in this paper, the received signal strength depends only on the difference in received signal strengths. Without loss of generality, we let $K_1 = 0dB$.

The log-normal shadowing is assumed to have exponential autocorrelation and independence between base stations. This independence is not a problem since we investigate only strategies based on the difference in received signal strengths. The exponential correlation approximation, based on experimental results [2], is

$$\begin{aligned} \mathrm{E}\{u(k_1 d_s)u(k_2 d_s)\} &= \mathrm{E}\{v(k_1 d_s)v(k_2 d_s)\} \\ &= \sigma_s^2 \exp\left(\frac{-d_s|k_1-k_2|}{d_o}\right), \end{aligned} \tag{2}$$

where d_o, the *correlation distance*, determines how fast the correlation decays with distance. Unless noted otherwise, we assume that $d_o = 20m$ as was done in [6, 8, 9]. The measurements are averaged using an exponential averaging window where the constant d_{av}, the *averaging distance*, determines how much weight is given to past values of the signal. The discrete-time counterpart of exponential smoothing has a simple recursive form for the measured signals from base A and B given as

$$\begin{aligned} \bar{a}_k &= \mathrm{e}^{-d_s/d_{av}}\, \bar{a}_{k-1} + \left(1 - \mathrm{e}^{-d_s/d_{av}}\right) a_k, \\ \bar{b}_k &= \mathrm{e}^{-d_s/d_{av}}\, \bar{b}_{k-1} + \left(1 - \mathrm{e}^{-d_s/d_{av}}\right) b_k. \end{aligned} \tag{3}$$

In this paper we consider handover algorithms based on *relative signal strength with hysteresis* with no threshold. The instantaneous and averaged signal strength differences are defined as

$$x_k = a_k - b_k \qquad \text{and} \qquad \bar{x}_k = \bar{a}_k - \bar{b}_k. \tag{4}$$

Since a_k and b_k are Gaussian random variables, x_k and $\bar{x}_k$ are also Gaussian. [1]

Let $A(k-1)$ denote the event that the serving base station is A at $(k-1)$st sampling instant. Then $P_{B|A}(k)$, the probability of handing over from base A to B at sampling instant k, and $P_{A|B}(k)$, the probability of handing over from base B to A at k, are written [8] as

$$\begin{aligned} P_{B|A}(k) &= P\{\bar{x}_k < -h \mid A(k-1)\}, \\ P_{A|B}(k) &= P\{\bar{x}_k > h \mid B(k-1)\}. \end{aligned} \tag{5}$$

The probability of handover at k is then given as

$$P_{ho}(k) = P_{B|A}(k)P_A(k-1) + P_{A|B}(k)P_B(k-1). \tag{6}$$

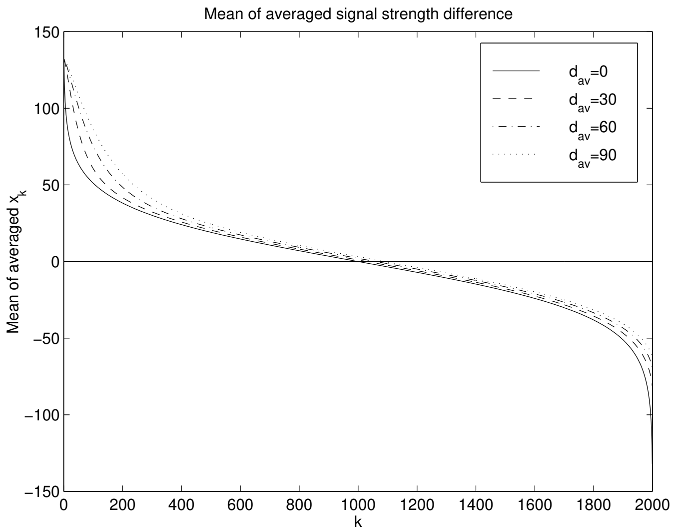

Figure 2: Mean of averaged signal strength, $\bar{m}_k$ for averaging distance $d_{av} = 30, 60$ and 90m.

The handover delay is defined in terms of the assignment probabilities of the mobile to both base stations; the handover delay $\hat{k}$ is the instant (potentially between sampling points) where the assignment probability of the mobile to base A is equal to 0.5, i.e., $P_A(\hat{k}) = P_B(\hat{k}) = 0.5$. The assignment probabilities of the mobile are given as

$$\begin{aligned} P_A(k) &= \left(1 - P_{B|A}(k)\right) P_A(k-1) \\ &\quad + P_{A|B}(k) P_B(k-1), \\ P_B(k) &= P_{B|A}(k) P_A(k-1) \\ &\quad + \left(1 - P_{A|B}(k)\right) P_B(k-1). \end{aligned} \tag{7}$$

The following sections evaluate the handover delay for the cases without and then with hysteresis.

3 Handover Delay without Hysteresis

This section investigates the handover delay for a system without hysteresis, $h = 0$. Recent work has proposed a delay expression for the case of independent log-normal shadowing and fixed duration sliding window averaging [7].

For a system without hysteresis, this section provides an exact analytic solution, that includes the correlation of shadow fading and the more commonly used exponential averaging. In addition, it provides a simple algebraic approximation that works well over the range of interest.

Without hysteresis, the handover decision is memoryless, i.e. the probability of being assigned to a certain base does not depend on the base it was assigned to in the past. The probability of being assigned to base A at time k is

$$P_A(k) = P\{\bar{x}_k > 0\} = \mathrm{Q}\left(\bar{m}_k / \bar{\sigma}_k\right), \tag{8}$$

where $\bar{x}_k$ is a Gaussian random variable with mean $\bar{m}_k$ and variance $\bar{\sigma}_k^2$. Since the Q-function[1] attains value 1/2 when its argument is equal to zero, the handover delay, $\hat{k}$, is found as the solution of

$$\bar{m}_k = 0 \tag{9}$$

From (4) the mean of x_k, m_k, is given as

$$m_k = K_2 \log\left(\frac{D-k}{k}\right). \tag{10}$$

From the definition of averaging given by (3), and the definition of $\bar{x}_k$ given by (4), it is possible to derive a closed form expression for the mean of $\bar{x}_k$ as

$$\bar{m}_k = \left(1 - e^{-d_s/d_{av}}\right) \sum_{j=0}^{k} e^{-j d_s/d_{av}} m_{k-j}. \tag{11}$$

It can be shown by using a series expansion around $D/2$ that m_k is linear around $D/2$. Figure 2 shows $\bar{m}_k$ for $d_{av} = 0, 30, 60, 90$m for $D = 2000$m and $K_2 = 40$ dB. Note that it is linear for a significant portion of the distance between the base stations. The effect of exponential averaging on an infinitely long linear function is to shift it to the right by an amount equal to d_{av}, the averaging distance. In other words, in the linear region, $\bar{m}_k \approx m_{k-d_{av}}$, where the relation is an approximation for two reasons: the curve only approximately linear over a limited region, and the curve is of finite length. Since $m_{D/2} = 0$, the solution of (9), $\hat{k}$, is given by

$$\hat{k} \approx \frac{D}{2} + d_{av}. \tag{12}$$

This result has been confirmed by numerical evaluation of (11), shown as asterisks on Figure 3. The line is the linear approximation given by (12). The approximation and the exact solution are in excellent agreement over the range of interest.

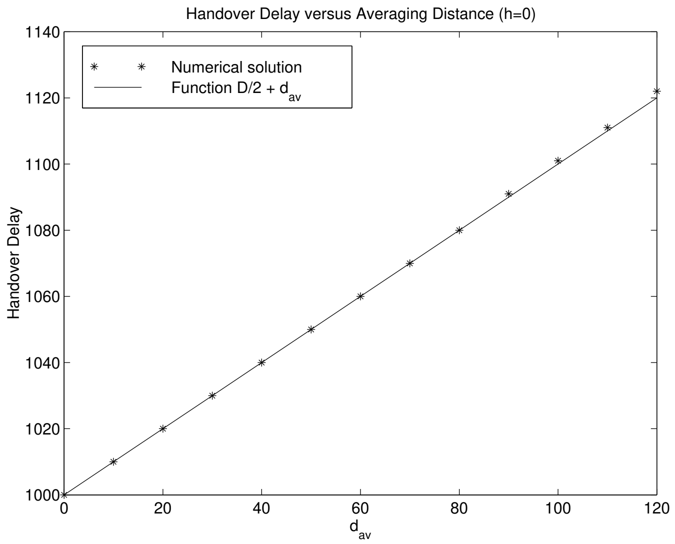

Figure 3: Exact (numerical) and approximate results for handover delay versus averaging distance.

[1] The "Q" function is defined as $\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy$.

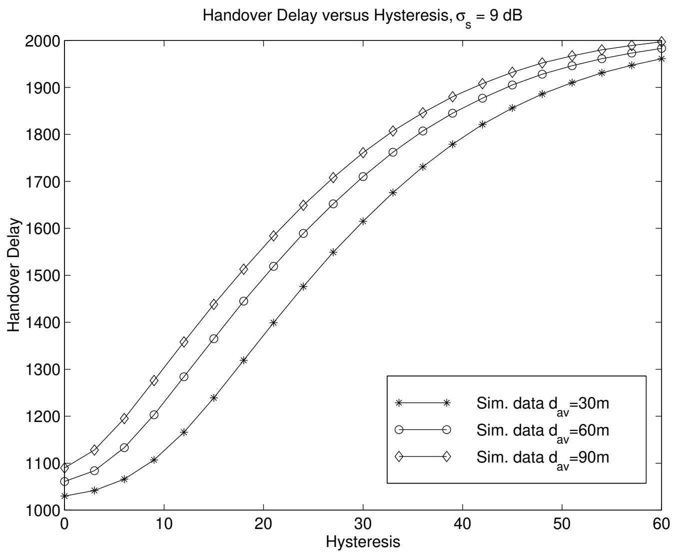

Figure 4: Handover delay versus hysteresis generated by simulation for $\sigma_s = 9$ dB, $d_{av} = 30, 60$ and 90m.

4 Handover Delay with Hysteresis

This section extends the result of the previous section to include a hysteresis margin. Finding an exact analytic expression for $P_A(k)$ does not seem tractable. For this reason we will use data generated by simulation and apply curve fitting techniques to come up with an approximate closed form expression for the handover delay. Figure 4 shows the handover delay as a function of hysteresis for three values of averaging $d_{av} = 30, 60$ and 90m, when all other system parameters are fixed: $D = 2000$m, $K_2 = 40$dB, $\sigma_s = 9$dB, $d_o = 20$m. As expected the handover delay increases with increasing hysteresis, h, and averaging distance, d_{av}. By definition, handover delay is larger than $D/2$. Thus, it is convenient to separate $D/2$ from the whole delay expression. The S-shaped curves with well-defined asymptotes seen in Figure 4 (starting at $D/2 + d_{av}$ and ending at D) suggest the use of a function of the form[2]

$$\hat{k} = f(h) = \frac{D}{2} + \frac{c_1}{1 + c_2 \exp(-\frac{h^a}{b})}, \tag{13}$$

shown in Figure 5.

Assume the maximum value handover delay can take is D, since it is the separation between the base stations. Equating the maximum value of $f(h)$, $D/2 + c_1$, to D ($\lim_{h\to\infty} f(h) = D/2 + c_1 = D$), c_1 is obtained as $D/2$. From the previous section we know that at $h = 0$ the handover delay is equal to $D/2 + d_{av}$. Equating $f(0)$ to $D/2 + d_{av}$ yields $c_2 = (D/2 - d_{av})/d_{av}$. Therefore, the function to be used approximating the handover delay becomes

$$\hat{k} = \frac{D}{2} + \frac{D/2}{1 + \frac{D/2 - d_{av}}{d_{av}} \exp(-\frac{h^a}{b})}. \tag{14}$$

[2] This function is commonly used to model the penetration for a new service by economists, and is often called a Fisher-Pry model [3]. There is no physical justification for its use in this application other than its simplicity and, in hindsight, goodness of fit.

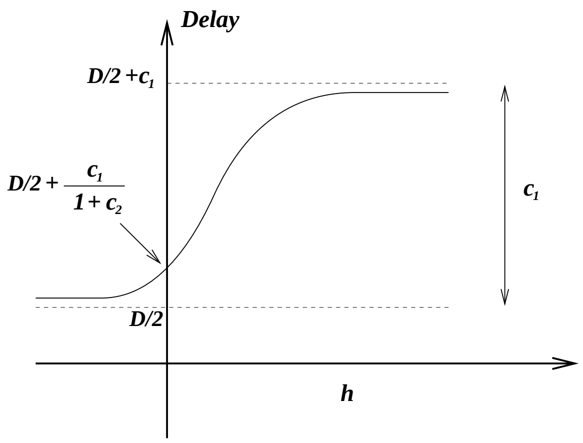

Figure 5: Functional form to be used in approximating the handover delay.

As seen from the example for $\sigma_s = 9$ dB in Figure 4, for large hysteresis, the handover delay may become undesirable large. In a real cellular wireless system, a handover delay more than 25% of the distance between midpoint (at $D/2$) and the target base station (at D) is not acceptable.[3] Therefore, we will concentrate on hysteresis levels and averaging distance values which will yield handover delays between $D/2$ and $1.25(D/2)$ in order to get more accurate approximations.

This paper assumes that the path loss exponent is fixed and equal to $K_2 = 40$ dB. The parameters a and b in (14) may depend on σ_s, the standard deviation of the log-normal shadowing, and d_{av}, the averaging distance. Through an extensive nonlinear curve fitting study we observed that a depends only on σ_s and b depends on σ_s and d_{av}. For $K_2 = 40$ dB, a and b have the following forms:

$$\begin{aligned} a &= a_1 \sigma_s^{a_2}, \\ b &= b_1 \sigma_s^{b_2} + b_3. \end{aligned} \tag{15}$$

where $a_1 = 0.55$, $a_2 = 0.32$ and $b_1 = 0.03$ independent of d_{av}. On the other hand b_2 and b_3 depend on d_{av}. In particular b_2 decreases and b_3 increases linearly with d_{av}. The dependence of a and b on the path loss exponent has been investigated as well, but will not be presented here due to space limitations.

In Figures 6 and 7, the approximate closed form expression found for handover delay is plotted along with simulation data for $\sigma_s = 6$ dB and 12 dB, respectively, for averaging distance $d_{av} = 30, 60$ and 90 meters. In these figures, solid, dashed and dotted curves show the function as a continuos function of h and the symbols (star, circle and diamond) show the simulation data points.

[3] Assuming a TDMA system. In a CDMA system with soft handover, this is acceptable and often desirable.

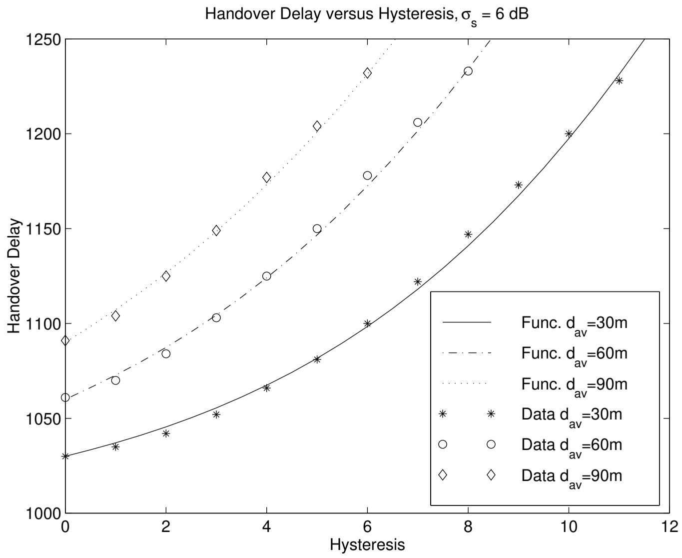

Figure 6: Closed form expressions and simulation data plotted for $\sigma_s = 6$ dB and $d_{av} = 30, 60$ and 90m.

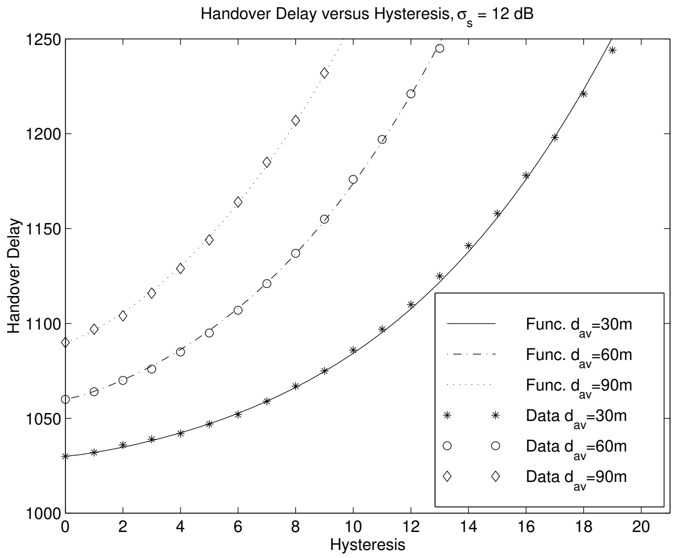

Figure 7: Closed form expressions and simulation data plotted for $\sigma_s = 12$ dB and $d_{av} = 30, 60$ and 90m.

We observe that the approximation works well over the range of interest. Similar results were obtained for other values of the path loss exponent.

5 Summary

The performance of a handover algorithm is often measured in terms of two criteria: expected number of handovers and handover delay. The expected number of handovers and handover delay are inversely proportional in the sense that increasing the hysteresis margin and/or averaging distance decreases the expected number of handovers but increases the handover delay. The exact tradeoff, however, has not been determined except through case-by-case simulation studies and very complex analytic approximations. Given a propagation model and fixed system parameters it is desirable to find best values of design parameters, the averaging distance and the hysteresis margin, that will yield acceptable values in terms of these two performance metrics.

This paper has developed a simple closed form approximation for handover delay in terms of base station separation (D), hysteresis margin (h), averaging distance (d_{av}), and the standard deviation of log-normal shadowing (σ_s) for a fixed path loss exponent $K_2 = 40dB$, and a fixed correlation distance $d_o = 20m$. Other path loss exponents have been studied elsewhere. Those results indicate that the approximation works well for general K_2.

The approximate closed form expression developed in this paper for handover delay along with the one developed in [5] for the expected number of handovers can be combined to study the trade-off between the expected number of handovers and handover delay. It is the topic of ongoing work.

References

[1] W. B. Davenport, Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise,* McGraw Hill, 1958; New York: IEEE Press, 1987.

[2] M. Gudmundson, "Analysis of Handover Algorithms," *Proc. Vehicular Technology Conf. 91,* St. Louis, MO, pp. 537-542, May 19-22, 1991.

[3] J. Martino, *Technology Forecasting for Decision Making,* McGraw-Hill, 3rd Edition, 1993.

[4] G. P. Pollini, "Trends in Handover Design," *IEEE Comm. Mag.,* Vol. 34, no. 3, pp. 82-90, March 1996.

[5] G. P. Pollini, "Handover Rates in Cellular Systems: Towards a Closed Form Approximation," *Proc. Globecom 97,* Phoenix, AZ, pp. 711-715, Nov. 2-7, 1997.

[6] R. Vijayan, and J. M. Holtzman, "A model for Analyzing Handoff Algorithms," *IEEE Trans. Vehicular Tech.,* Vol. 42, no. 3, pp. 351-356, Aug. 1993.

[7] M. Zonoozi, and P. Dassanayake, "Handover Delay and Hysteresis Margin in Microcells and Macrocells," *Proc. PIMRC 97,* Helsinki, Finland, pp. 396-400, Sept 1-4, 1997.

[8] N. Zhang, and J. M. Holtzman, "Analysis of Handoff Algorithms Using Both Absolute and Relative Measurements," *Proc. Vehicular Technology Conf. 94,* Stockholm, Sweden, pp. 82-86, Jun 8-10, 1994.

[9] N. Zhang, and J. M. Holtzman, "Analysis of Handoff Algorithms Using Both Absolute and Relative Measurements," *IEEE Trans. Vehicular Tech.,* Vol. 45, no. 1, pp. 174-179, Feb. 1996.