Minimum-Risk Layer 2 Trigger Levels for Proactive Media-Independent Handovers

David Griffith, Alexandre Delye de Clauzade de Mazieux, and Nada Golmie
National Institute of Standards and Technology (NIST), Gaithersburg, MD, 20899 USA
Email: david.griffith@nist.gov (corresponding author)

Abstract—Anticipated handovers that use Link Going Down (LGD) and Link Down (LD) trigger events require the network operator to set the LGD trigger high enough that the handover completes before the LD trigger event. However, setting the LGD trigger too high can result in frequent handovers by mobile nodes, leading to high signaling overhead. We propose a mechanism for balancing these requirements using a risk function. The function expresses the risk with respect to the probability that the LD event falls within a range of times with respect to handover completion. The risk function can be used with weights that allow the network operator to set the relative importance of the early handover completion requirement and the requirement that the LGD trigger not be too sensitive. Because the risk is expressed using the characteristic function of the handover time, we can easily use it to set the LGD trigger for any mobility management protocol.

Index Terms—anticipated handovers, IEEE 802.21, optimization, performance modeling, mobility models

I. INTRODUCTION

Proactive handovers allow users to be connected to voice, data, and video services while remaining mobile. The work by the IEEE 802.21 Media Independent Handover (MIH) group uses channel state reports from the link layer to generate event triggers that enable MIH application software in a migrating mobile to start setting up a new connection at the access network the mobile is entering while the old connection is still viable [1]. When MIH triggers are used in conjunction with a mobility management protocol such as Fast Mobile Internet Protocol (FMIP), it becomes possible to create a tunnel between the routers at the mobile’s current access network and the new access network, so that packets can be forwarded to the mobile at its new location, even before the mobile’s home agent creates a fresh binding update.

The MIH event triggers are activated by changes in channel state variables such as the received signal strength (RSS). When the RSS falls below a certain threshold, a Link Going Down (LGD) trigger fires; this acts as a warning that loss of signal will occur soon. A fall in RSS below a lower threshold results in a Link Down (LD) trigger firing; it indicates that the RSS is not sufficient to support communication between the mobile and its access point (AP) [2].

Setting a threshold value for the RSS that causes the LGD trigger to fire involves a tradeoff between two competing design constraints. A high threshold will start the handover process well before the mobile’s connection breaks, thus minimizing the probability that the user’s connections are disrupted. However, early handover requires reserving resources (e.g., bandwidth for tunnels) that may be in short supply; a network operator will not want to allow users to hold these extra resources for too long. Conversely, setting the LGD trigger threshold low reduces the probability that the handover will take place too soon but allows less time to complete the handover and increases the chances that the link to the mobile will go down before the handover completes, resulting in dropped packets.

Previous work has examined an approach in which the Link Going Down trigger is estimated using predictive methods such as Least Mean Square (LMS) estimation or, more simply, an estimator based on the slope of the declining received signal strength [3]. Our approach is complementary to this technique and could be used jointly with it, by setting initial values for the triggers that could be modified in response to changing channel conditions.

In this paper, we compute the probabilities of the two undesirable events described above, and use them to develop a weighted cost function that encapsulates the risk associated with a particular value of the LGD trigger. Rather than use the LGD trigger value directly, we use the mean time from the LGD trigger to the Link Down event. The mean time can be mapped to a LGD trigger value by accounting for the mobile’s velocity as well as channel conditions. By varying the weights in the cost function, the network operator can tailor a LGD trigger value based on factors such as QoS and network resource availability. An additional feature of our approach is that, by conservatively modeling the time from LGD to LD with an exponential distribution, we can express the risk using the characteristic function of the time to complete the handover. In this paper, we use a shifted gamma distribution to model the handover time, but customized models can be used as well.

The remainder of this paper is organized as follows. In Section II, we show that the time between the LGD and LD trigger events has a gamma distribution, and that it can
be modeled with an exponential distribution if the trigger thresholds are close. In Section III, we use the inter-trigger time model to derive a risk function and we show how to solve it to get the optimal LGD threshold. In Section IV, we plot the optimal handover operating point with respect to various parameter values. We also use simulations to evaluate our approach using a two dimensional random walk for the mobility model. In Section V, we use packet level simulations to show how one can use measured network statistics to construct a risk function and use it to find an optimal value for the LGD trigger. We summarize our results in Section VI.

II. MODELING LGD-LD INTER-TRIGGER DURATION

We define $X$ to be the time between the LGD and LD triggers. Its distribution depends on the movement of the mobile node. The amount of time that a mobile spends inside the coverage area of an AP has been the subject of considerable research. Early work focused on the call duration time, which was initially assumed to be exponentially distributed [4], which was soon confirmed by using the gamma cumulative distribution function. The dwell time of a mobile was shown via computer simulation using a novel mobility model to have a gamma distribution by Zonoozi and Dassanayake [6]. This work showed that the gamma distribution was a good fit for mobile nodes whose call duration time, which was initially assumed to be exponentially distributed [4], was soon confirmed by using the gamma cumulative distribution function. The dwell time of a mobile was shown via computer simulation using a novel mobility model to have a gamma distribution. To compute the chi square test statistic:

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{(E_i - O_i)^2}{E_i},$$

where $E_i$ and $O_i$ are respectively the expected and observed counts for the $i$th bin. We obtain the values of $E_i$ by using the gamma cumulative distribution function. Assuming none of the bins is empty (i.e. has no sample values), the number of degrees of freedom for the test is $(N_{\text{bins}} - 1) - 3$, since we have to estimate three parameters for the model distribution: the shape parameter $a$, the offset $b$, and the scale parameter $\lambda$.

We estimate the distribution parameters as follows. We can get the estimate for $b$ directly, which greatly simplifies the computations for obtaining estimates for $a$ and $\lambda$. Given that the mobile’s velocity $v$ is constant, it follows that $b$ is the minimum time for the mobile to go from $R_{\text{LGD}}$ to $R_{\text{LD}}$, which it does by moving on a radial line. Thus

$$b = \frac{R_{\text{LD}} - R_{\text{LGD}}}{v}. \tag{1}$$

To get estimates for the shape and scale parameters $a$ and $\lambda$, we use Fisher’s maximum likelihood technique. The natural logarithm of the likelihood function for a vector $x = [x_1, x_2, \ldots, x_n]$ of $n$ samples from a gamma distribution is

$$L(a, \lambda; x) = (a - 1) \sum_{i=1}^{n} \ln(x_i - b) - \frac{1}{\lambda} \sum_{i=1}^{n} (x_i - b) - n \cdot a \ln(\lambda) - n \ln(\Gamma(a)),$$

where $x_i \geq b$ for $i = 1, 2, \ldots, n$. The values of $\hat{a}$ and $\hat{\lambda}$ that maximize the likelihood function are those that satisfy the following equations, respectively:

$$\frac{\partial}{\partial a} L(a, \lambda; x) = \sum_{i=1}^{n} \ln(x_i - b) - n \ln(\lambda) - n \psi^{(0)}(a) = 0 \tag{3}$$

$$\frac{\partial}{\partial \lambda} L(a, \lambda; x) = \frac{1}{\lambda^2} \sum_{i=1}^{n} (x_i - b) - \frac{n \cdot a}{\lambda} = 0 \tag{4}$$

where

$$\psi^{(k)}(a) = \frac{d^k}{da^k} \ln(\Gamma(a)), \quad k = 0, 1, \ldots$$

is the polygamma function, with $k = 0$ in this case. By solving Eq. (4) for $\lambda$ and substituting the result into Eq. (3), we get the following expression:

$$\ln(a) - \psi^{(0)}(a) + \ln(G(x - b)/A(x - b)) = 0, \tag{5}$$

where

$$G(x - b) = \left( \prod_{i=1}^{n} (x_i - b) \right)^{1/n} \tag{6}$$

and

$$A(x - b) = \frac{1}{b} \sum_{i=1}^{n} (x_i - b) \tag{7}$$
are, respectively, the geometric and arithmetic means of the data set \( x \) shifted by \( b \).

Defining \( f(a) = \ln(a) - \psi^{(0)}(a) + \ln(G(x - b)/A(x - b)) \), we use Newton’s tangent method to find the maximum likelihood estimator for \( a \) by finding the zero of \( f(a) \). Our starting point is the estimate of \( a \) that results from using the method of moments [8]. The first two raw moments of \( X \), \( m_1 = \mathbb{E}(X) \) and \( m_2 = \mathbb{E}(X^2) \), are:

\[
\begin{align*}
  m_1 &= b + a\lambda \\
  m_2 &= b^2 + 2ab\lambda + a(a + 1)\lambda^2.
\end{align*}
\]

Solving the system of equations in Eq. (8)–(9) gives us two estimators, of which we need only \( \hat{a} \) and \( \hat{\lambda} \):

\[
\hat{a} = \frac{(m_1 - b)^2}{m_2 - m_1^2} \quad \text{and} \quad \hat{\lambda} = \frac{m_2 - m_1^2}{m_1 - b}.
\]

We compute \( \hat{a} \) by using the \( n \) samples in \( x \) to get sample values of \( m_1 \) and \( m_2 \). Starting with an initial value for the maximum likelihood estimate \( \hat{a}_0 = \bar{a} \), we generate successive estimates \( \hat{a}_i \) as follows:

\[
\begin{align*}
  \hat{a}_{i+1} &= \hat{a}_i - \frac{f(\hat{a}_i)}{f'(\hat{a}_i)} \\
  &= \hat{a}_i - \frac{\ln(\hat{a}_i) - \psi^{(0)}(\hat{a}_i) + \ln(G(x - b)/A(x - b))}{\hat{a}_i^2 - \psi^{(1)}(\hat{a}_i)}.
\end{align*}
\]

until the change in the estimator, \( \delta \hat{a}(i + 1) = |\hat{a}_{i+1} - \hat{a}_i| \), decreases below a given threshold, which we chose to be \( 10^{-6} \) for our computations. Once we have \( \hat{a} \), we use Eq. (4) to get the maximum likelihood estimator for \( \lambda \), \( \hat{\lambda} = A(x - b)/\bar{a} \).

In Figs. 2 and 3 we plot results from a series of Monte Carlo simulations in which we modeled a circular coverage area with a LD boundary located at a distance \( R_{\text{LGD}} = 100 \) m from the AP. We let the value of the distance from the AP to the LGD boundary, \( R_{\text{LGD}} \), vary from 97 m to 99 m. For each value of \( R_{\text{LGD}} \), we performed 500 trials; each trial consisted of 50 random walks. Each random walk was performed by a single mobile node traveling at a speed of \( v = 1 \) m/s, with position updates occurring every 1 s. The mobile node began each walk on the LGD boundary, with the phase of its location uniformly distributed over \([0, 2\pi]\), and its original direction of movement \( \theta_0 \) equal to the phase of its location, so that it was originally oriented along a radius of the LD circular boundary. At the \( k \)th position update, the mobile node updated its direction \( \theta_k \) by adding a random angle as follows: \( \theta_{k+1} = \theta_k + \phi_k \), where \( \phi_k \) was uniformly distributed over the range \([-3\pi/5, 3\pi/5]\).

We show the effect of restricting the value taken by \( \phi_k \) to this range in Fig. 1, which compares sample paths of two-dimensional random walks where the update angle \( \phi_k \) is uniform over the whole unit circle versus the case where it is uniform over a wedge defined by the range \([-3\pi/5, 3\pi/5]\). For both cases, the mobile node had a speed of 2 m/s and updated its position every 1 s. When \( \phi_k \sim U[-\pi, \pi] \), the mobile’s path tends to exhibit behavior closer to Brownian motion, including backtracking, as shown in Fig. 1(a). In contrast, restricting the range of motion of the mobile by letting \( \phi_k \sim U[-3\pi/5, 3\pi/5] \) results in a much more direct path from the LGD boundary to the LD boundary as shown in Fig. 1(b). This reduces the mean time between LGD and LD trigger events.

Fig. 2 shows the results of the chi square goodness of fit test examining fit of random walk data to gamma distributions, versus \( R_{\text{LGD}} \), the radius of the LGD boundary.
close together.

Using these results, we can obtain the distance between \(R_{\text{LD}}\) and \(R_{\text{LGD}}\) as a function of the expected value of \(X\). With this information, we can obtain the value of \(R_{\text{LGD}}^*\) from \(\mu_X\). For the parameters that we used in generating Figs. 2 and 3, we plot \(\mu_X\) versus \(R_{\text{LGD}}\) in Fig. 4. To generate each point in these plots, we used 500 Monte Carlo simulations, each of which consisted of 50 random walks, where \(\phi_k \sim U[-3\pi/5, 3\pi/5]\). For each set of random walks, we computed the maximum likelihood estimators of the \(X\)'s distribution parameters \(\hat{a}\) and \(\hat{\lambda}\), along with \(\hat{b}\) from Eq. (1), and obtained \(\hat{\mu}_X = \hat{b} + \hat{a}\hat{\lambda}\) for that set of data.

The resulting \(\hat{\mu}_X\) values are plotted in Fig. 4; each plotted point is located at the average value of \(\hat{\mu}_X\), with error bars that indicate deviations of one standard deviation from the mean. Note that the relationship between \(\hat{\mu}_X\) and \(R_{\text{LGD}}\) is nearly linear, even when \(R_{\text{LGD}}\) is close to \(R_{\text{LD}}\), as shown in the inset for Fig. 4. This linear dependence means that, if the mobile’s motion can be modeled, it is easy to obtain \(R_{\text{LGD}}^*\) from \(\mu_X\). In this case, a linear curve fit to two significant digits yielded the following relationship:

\[
R_{\text{LGD}} = 100 - \frac{\mu_X}{2}.
\]  

This result is appealing because it gives us \(\mu_X = 0\) m when \(R_{\text{LGD}} = 100\) m, which is what we expect.

We also examined the case where \(\phi_k \sim U[-\pi, \pi]\); the resulting plot of \(\hat{\mu}_X\) versus \(R_{\text{LGD}}\) appears in Fig. 5. Here we examined only values of \(R_{\text{LGD}}\) between 96 m and 100 m, and we performed 1000 Monte Carlo trials per data point, with each trial consisting of 50 random walks. While the mean and standard deviation of the results are both much greater than in the case where \(\phi_k \sim U[-3\pi/5, 3\pi/5]\), we nevertheless see the same linear relationship between \(\mu_X\) and \(R_{\text{LGD}}\), which indicates that the linear relationship holds even if we restrict the range of values that the mobile’s update angle can take. In fact, performing a linear curve fit to two significant digits, using the data in Fig. 5, gives

\[
R_{\text{LGD}} = 100 - \frac{\mu_X}{200},
\]

showing the effect that reducing the range of \(\phi_k\), by \(4\pi/5\) in this case, has on the slope of the line.

Once we have \(R_{\text{LGD}}\), we can easily map it to a RSS threshold for the Link Going Down trigger. For example, if there are no obstructions or sources of reflections, we can use the free space path loss model [9]:

\[
L_{\text{path}} = -20 \log_{10}(R_{\text{LGD}}) - 20 \log_{10}(f) - 20 \log_{10}(4\pi/c_{\text{air}}),
\]

where \(f\) is the center frequency of the radio signal and \(c_{\text{air}} = c/n_{\text{air}}\) is the speed of light in air, where \(n_{\text{air}} \approx 1.00029\) is air’s index of refraction at standard temperature and pressure. If we consider an example case where we have a IEEE 802.11b WiFi AP whose transmit power is 20 dBm and whose antenna has a gain of 4 dBi and a receiver whose receive gain is 2 dBi, and we are using Channel 9, whose center frequency is 2.452 GHz,
Handover Begins Link Down
Handover Complete Link Down
acceptable range for X

Fig. 6. Timeline of Link Going Down event and Link Down event, with handover beginning at time \( t = 0 \) and completion occurring at time \( t = H \).

a \( R_{\text{LGD}} \) value of 95 m corresponds to a LGD threshold of \(-53.8\) dBm.

III. HANDOVER RISK FUNCTION

There are two metrics that we will use to assess the performance of the handover scheme. These are \( P_D \), the probability that the existing wireless link breaks before the handover completes; and \( P_T \), the probability that the handover causes network resources to be reserved for longer than some maximum allowable time. By using the weighted sum of these probabilities, we will construct a risk function that we can use to find the optimal expected time from the generation of a Link Going Down (LGD) trigger to the generation of the Link Down (LD) trigger. We can use this optimal value to assign a threshold value for the RSS at which the LGD trigger will fire.

A. Performance metrics \( P_D \) and \( P_T \)

In Fig. 6, we show the sequence of major events associated with an anticipated handover based on link layer triggers. We define the time when the LGD trigger occurs to be \( t = 0 \). Let \( H \) be the amount of time to perform the handover setup procedure, and let \( X \) be the amount of time from LGD to LD, as shown in Fig. 6 for the case where the handover completes before the LD event. Both \( H \) and \( X \) are random variables with respective distributions \( F_H \) and \( F_X \). The probability that the link goes down before the handover setup completes is

\[
P_D = \Pr\{X \geq H\} = \int_0^\infty \Pr\{X \leq u\} f_H(u) \, du
\]

This assumes that the density \( f_H \) exists; if not, we can write \( P_D \) using the Lebesgue-Stieltjes integral \( P_D = \int_0^\infty F_X(u) \, dF_H(u) \). Alternatively, we can write \( P_D \) in terms of the density of \( X \) and the cumulative distribution of \( H \) as follows:

\[
P_D = \int_0^\infty [1 - F_H(u)] f_X(u) \, du,
\]

using integration by parts in Eq. (12) or the fact that \( P_D = 1 - \Pr\{H < X\} \).

If the LGD trigger is set high, the handover will complete well before the LD trigger fires. Thus, an excessively sensitive LGD threshold can cause handovers while a mobile is still well inside the coverage area of its previous AP. It can also increase the number of arrivals into destination access networks or cause ping-ponging between access networks if there are areas of deep fading within the coverage area. To reduce the risk of premature handovers, we define \( P_T \) to be the probability that the time when the LD trigger occurs, \( t = X \), is less than the handover completion time \( H \) plus some maximum tolerable amount of time \( \gamma \). We have

\[
P_T = \Pr\{X \leq H + \gamma\} = \int_0^\infty F_X(u + \gamma) f_H(u) \, du.
\]

We can get an alternative expression for \( P_T \), using the assumption that \( H \) and \( X \) are independent, by integrating over the appropriate region under the joint density of \( H \) and \( X \):

\[
P_T = \int_0^\infty \int_{u - \gamma}^{\infty} f_X(x) f_H(h) \, dx \, dh
\]

\[
= 1 - \int_0^\infty F_H(x - \gamma) f_X(x) \, dx.
\]

If we assume that \( X \) is exponentially distributed, using our results from Sec. II, we can develop a the following pair of equations for \( P_D \) and \( P_T \). Both equations are functions of the characteristic function \( \Phi_H(u) = \int_0^\infty f_H(x)e^{-ju} \, dx \). We first use Eq. (14) to compute \( P_T \):

\[
P_T = \int_0^\infty (1 - e^{-\left(u+\gamma\right)/\mu X}) f_H(u) \, du
\]

\[
= 1 - e^{-\gamma/\mu X} \int_0^\infty f_H(u)e^{-j\mu X} \, du
\]

\[
= 1 - e^{-\gamma/\mu X} \Phi_H(-j/\mu X).
\]

Setting \( \gamma = 0 \), we get the expression for \( P_D \):

\[
P_D = 1 - \Phi_H(-j/\mu X).
\]

For a given value of the tolerance, \( \gamma \), we can characterize the performance of a handover scheme using a particular LGD trigger threshold by plotting \( P_T \) versus \( P_D \). Plots of this type are a well-known tool for evaluating the performance of a system that must balance two competing goals; a well-known example is the radar receiver operating characteristic shows a performance curve in which we plot \( P_T \) versus \( P_D \); we wish to maximize and to minimize these quantities, respectively.

In Fig. 7, we show a set of operating characteristics for four different expected handover durations. All four sets of operating characteristics assume that the handover duration has a gamma distribution with shape parameter \( \alpha_H = 3 \), and offset \( b_H = 0.2 \) s. In each plot we considered four values of the tolerance parameter \( \gamma \). The plots show that the handover performance improves as the average

© 2011 ACADEMY PUBLISHER
handover time decreases, and that the performance becomes closer to ideal as we relax the tolerance. Indeed, a tight tolerance coupled with a large mean handover duration produce an operating characteristic that is close to being the worst possible one, as shown in Fig. 7(d). As the mean handover duration becomes large, we also see that the sensitivity of the operating characteristic to the tolerance variable \( \gamma \) decreases. This indicates that we may be forced to accept a large acceptable range of times in which the Link Down trigger can occur if the mean handover time is particularly large.

### B. The Risk Function

In order to optimize the handover performance we introduce a risk function \( R(\mu_X) \) that includes the weighted risks of the two events that we considered above: loss of signal prior to handover completion; and initiating the handover too soon. The risk is a function of the mean handover time, \( \mu_X \), because \( \mu_X \) depends on the threshold that the network operator chooses for the LGD event. The risk is

\[
R(\mu_X) = C_D \text{Pr}\{X \leq H\} + C_T \text{Pr}\{X > H + \gamma\} = C_D P_D + C_T (1 - P_T),
\]

where \( C_D \) and \( C_T \) are the costs associated with the events \( \{X \leq H\} \) and \( \{X > H + \gamma\} \), respectively. Using Eq. (16) and Eq. (17), we get the risk in terms of the characteristic function of the handover duration, \( \Phi_H(\omega) \):

\[
R(\mu_X) = C_D - (C_D - C_T e^{-\gamma/\mu_X}) \Phi_H(-j/\mu_X). \tag{19}
\]

Note that as the mean time from the LGD event to the LD event becomes large, the cost associated with handing over too soon dominates the risk: \( \lim_{\mu_X \to \infty} R(\mu_X) = C_T \). Likewise, if the mean time between the LGD and LD events is very small, the cost associated with a premature loss of signal dominates, i.e. \( \lim_{\mu_X \to 0} R(\mu_X) = C_D \).

We can minimize \( R(\mu_X) \) by taking the derivative with respect to \( \mu_X \), which gives us

\[
\frac{dR(\mu_X)}{d\mu_X} = \frac{1}{\mu_X} \left( C_T e^{-\gamma/\mu_X} \Phi_H(-j/\mu_X) + j \left( C_D - C_T e^{-\gamma/\mu_X} \right) \Phi'_H(-j/\mu_X) \right).
\]

(20)

The risk function has a minimum when \( dR(\mu_X)/d\mu_X = 0 \), meaning that the value of \( \mu_X \) that minimizes the risk satisfies the equation

\[
\frac{\Phi'_H(-j/\mu_X)}{\Phi_H(-j/\mu_X)} = \frac{j\gamma}{\mu_X}. \tag{21}
\]

In some cases, it may be easier to work with the logarithm of the characteristic function rather than the characteristic function itself. Since

\[
\frac{d}{d\mu_X} \ln(\Phi_H(-j/\mu_X)) = \frac{j\gamma}{\mu_X^2 \Phi_H(-j/\mu_X)},
\]

the minimization condition becomes

\[
\frac{d}{d\mu_X} \ln(\Phi_H(-j/\mu_X)) = \frac{j\gamma}{\mu_X^2 \Phi_H(-j/\mu_X)} \tag{22}
\]

We can develop a specific minimization condition by letting \( H \) have a shifted gamma distribution, which has the following form:

\[
F_H(x) = \begin{cases} 0, & x < b_H \\ \frac{1 - \Gamma((a_H/x - b_H)/\lambda_H)}{\Gamma(a_H)}, & x \geq b_H \end{cases} \tag{23}
\]

where \( \Gamma(w, x) \) is the (upper) incomplete gamma function having the form \( \Gamma(w, x) = \int_x^\infty e^{-u} u^{w-1} du \). The corresponding probability density function is

\[
f_H(x) = \begin{cases} 0, & x < b_H \\ \frac{(x-b_H)^{a_H - 1} e^{(x-b_H)/\lambda_H}}{\lambda_H^a \Gamma(a_H)}, & x \geq b_H \end{cases} \tag{24}
\]

The expected value of \( H \) is \( \mu_H = b_H + a_H \lambda_H \), and the variance is \( \sigma^2_H = a_H \lambda_H^2 \). The characteristic function is

\[
\Phi_H(\omega) = \frac{e^{-j\omega b_H}}{(1 + j\omega \lambda_H)^{a_H}} = \frac{a_H^{a_H} e^{-j\omega b_H}}{(a_H + j\omega(a_H - b_H))^{a_H}}. \tag{25}
\]

If we want to compute \( P_D \) and \( P_T \) for the case where \( H \) has a shifted gamma distribution and \( X \) has an exponential distribution, we can use direct computation via Eq. (12) or Eq. (13) and Eq. (14) or Eq. (15), respectively. An alternative and, in this case, easier, approach is to use the characteristic function in Eq. (25) in the expressions Eq. (17) and Eq. (16), which gives us

\[
P_D = 1 - \left( \frac{a_H \mu_X}{a_H \mu_X + \mu_H - b_H} \right)^{a_H} e^{-b_H/\mu_X} \tag{26}
\]

and

\[
P_T = 1 - \left( \frac{a_H \mu_X}{a_H \mu_X + \mu_H - b_H} \right)^{a_H} e^{-(b_H+\gamma)/\mu_X}. \tag{27}
\]
We can use the characteristic function from Eq. (25) to compute the risk function \( R(µ_X) \) and then take the derivative with respect to \( µ_X \). A more direct approach is to apply the characteristic function to Eq. (22); the resulting minimization equation is (after some simplification)

\[
\frac{a_H µ_X + b_H(µ_X - b_H)}{a_H µ_X + µ_X - b_H} = \frac{C_T γ}{C_T e^{γ/µ_X} - C_T},
\]

(28)

C. Existence of a Solution

To determine when a solution \( µ_X^* \) of Eq. (28) exists, it is helpful to examine the expressions on each side of the equality. The left-hand side, which we denote as \( l(µ_X) \), is a monotonically increasing function of \( µ_X \), since its derivative,

\[
\frac{dl(µ_X)}{dµ_X} = \frac{a_H(µ_X - b_H)^2}{(a_H µ_X + µ_X - b_H)^2},
\]

is greater than zero for all values of \( µ_X \). The derivative rapidly decays to 0 as \( µ_X \) increases due to the square in the denominator. Also, \( l(0) = b_H \) and \( \lim_{µ_X→∞} l(µ_X) = µ_H \). Because \( µ_H ≥ b_H \), the denominator of \( l(µ_X) \) never vanishes, so \( l(µ_X) \) is continuous for \( µ_X ∈ \{0, \infty\} \).

Examining the right-hand side of Eq. (28), which we denote as \( r(µ_X) \), \( r(0) = 0 \) and \( \lim_{µ_X→∞} r(µ_X) = C_T γ / (C_D - C_T) \).

If \( b_H = 0 \), then \( l(0) = r(0) = 0 \); but this is not a useful result, since setting \( µ_X = 0 \) is equivalent to setting the threshold for the LGD trigger to be so low that it is equal to that of the LD trigger.

If \( C_D ≤ C_T \), \( r(µ_X) \) has a vertical asymptote at \( µ_X^{max} = \gamma / \ln(C_T / C_D) \), and \( r(µ_X) < 0 \) when \( µ_X > µ_X^{max} \), and we can always find a value of \( µ_X \) that minimizes the cost function. This is because \( r(0) ≤ l(0) \) and \( \lim_{µ_X→µ_X^{max}} l(µ_X) < \lim_{µ_X→µ_X^{max}} r(µ_X) = ∞ \). Both \( l(µ_X) \) and \( r(µ_X) \) are continuous, so they intersect at some value \( 0 < µ_X^{max} ≤ µ_X^{max} \).

If \( C_D > C_T \), \( r(µ_X) \) is a continuous function of \( µ_X \) that is monotonically increasing because its derivative,

\[
\frac{dr(µ_X)}{dµ_X} = \frac{C_D C_T γ^2 e^{γ/µ_X}}{(C_D e^{γ/µ_X} - C_T)^2 µ_X^2},
\]

(29)

is positive and finite for all values of \( µ_X \). No solution to Eq. (28) exists if \( \frac{C_D C_T γ^2}{C_D - C_T} < b_H \), since \( \max_{µ_X} r(µ_X) < \min_{µ_X} l(µ_X) \). Conversely, a solution always exists if \( \frac{C_D C_T γ^2}{C_D - C_T} ≥ µ_H \), since \( r(0) ≤ l(0) \) and \( \lim_{µ_X→∞} r(µ_X) = \lim_{µ_X→∞} r(µ_X) = ∞ \), so \( l(µ_X) \) and \( r(µ_X) \) will intersect because both are continuous in \( µ_X \). For the case where \( b_H ≤ \frac{C_D C_T γ^2}{C_D - C_T} < µ_H \), a solution exists if \( r(µ_X) \) increases rapidly enough that it intersects \( l(µ_X) \). In such a situation, because \( \lim_{µ_X→∞} l(µ_X) > \lim_{µ_X→∞} r(µ_X), l(µ_X) \) and \( r(µ_X) \) must intersect at two values of \( µ_X \). Between these values, \( r(µ_X) > l(µ_X) \). Defining \( α = \frac{C_D}{C_T}, r(µ_X) \) increases most rapidly when \( α = 1 \). To show this, we write the derivative of \( r(µ_X) \) shown in Eq. (29) as a function of both \( µ_X \) and \( α \):

\[
δ_r(µ_X, α) = \frac{α γ^2 e^{γ/µ_X}}{(α e^{γ/µ_X} - 1)^2 µ_X^2},
\]

(30)

The partial derivative of Eq. (30) with respect to \( α \) is

\[
\frac{δ_r(µ_X, α)}{δα} = -\frac{γ^2 e^{γ/µ_X} (α e^{γ/µ_X} + 1)}{(α e^{γ/µ_X} - 1)^2 µ_X^2},
\]

which is negative for \( α ≥ 1 \) and \( µ_X ≥ 0 \), meaning that \( r(µ_X) \) increases most rapidly over the entire set of positive real numbers when \( C_D < C_T \).

We consider conditions for \( l(µ_X) \) and \( r(µ_X) \) to intersect when \( α = 1 \). If no intersection occurs for this value of \( α \), it will not happen when \( α > 1 \), since \( r(µ_X) \) is maximized for each value of \( µ_X \) when \( α = 1 \) and decreases with respect to \( α \). We consider two cases: \( \lim_{µ_X→∞} r(µ_X) = 0.50 µ_H \) and \( \lim_{µ_X→∞} r(µ_X) = 0.99 µ_H \), where the value is obtained by setting \( γ = k(µ_H - 1) \), where \( k \) is the desired limit value.

A larger value of \( \lim_{µ_X→∞} r(µ_X) \) with respect to \( µ_H = \lim_{µ_X→∞} l(µ_X) \) will cause \( r(µ_X) \) to increase to a greater terminal value and thus increase the range of other parameter values that will result in value of \( µ_X^* \) that satisfies the optimization condition. For each value of \( \lim_{µ_X→∞} r(µ_X) \), we let consider two values of \( µ_H \):

\( µ_H = b_H + 0.5 \) and \( µ_H = b_H + 5.0 \). A larger value of \( µ_H \) increases the range of values taken by \( l(µ_X) \).

For each set of values that we assigned to the parameters, we determined the largest value of \( H^i \)’s shift parameter \( b_H \) for which \( l(µ_X) \) and \( r(µ_X) \) intersected. The resulting curves are shown in Fig. 8. The curves show that decreasing \( \lim_{µ_X→∞} r(µ_X) \) reduces the range of \( b_H \) values for which an intersection occurs, given \( α_H \).

They also show that a similar reduction in the range of \( b_H \) values occurs when we reduce \( µ_H \). Most significantly, we observe that no intersection occurs when \( α_H ≥ 1 \), which corresponds to \( H \) having an exponential distribution. Because of the range of values for the other parameters that we considered, we can say that if \( C_D > C_T \) and \( α_H ≥ 1 \), then a solution to Eq. (28) exists only if \( \frac{C_D C_T γ^2}{C_D - C_T} > µ_H \).
D. A Simpler Risk Function

In this subsection, we consider an alternative risk function that is simpler in form because the costs associated with the two events $D$ and $T$ are equal. Recall that we are trying to ensure that the LD event does not occur before the handover completes while simultaneously trying to prevent the handover from occurring more than $\gamma$ s before the LD event. Thus we are seeking to maximize the probability of the event $\{H < X \leq H + \gamma\}$, which is

$$P_T\{H < X \leq H + \gamma\} = P_T - P_D = \Phi_H(-j/\mu_X) \left[1 - e^{-\gamma/\mu_X}\right],$$

(31)

again assuming that $X$ is exponentially distributed with mean $\mu_X$. The risk function we want to minimize is $1 - P_T\{H < X \leq H + \gamma\}$. If $H$ has the distribution in Eq. (23), the risk function is

$$R(\mu_X) = 1 - \frac{e^{-b_H/\mu_X} (1 - e^{-\gamma/\mu_X}) (a_H \mu_X)^{a_H}}{(a_H \mu_X + \mu_H - b_H)^{a_H}}.$$  

(32)

To find the minimum of this function of $\mu_X$, we take the derivative and set it to zero. The derivative is zero at $\mu_X = 0$, where the function has a maximum, and where $\mu_X$ satisfies the equation

$$a_H (\gamma + \mu_H) \mu_X + (b_H + \gamma) (\mu_H - b_H) = e^{\gamma/\mu_X} (a_H \mu_X b_H + \mu_H (\mu_H - b_H)).$$

(33)

IV. NUMERICAL RESULTS

In this section, we demonstrate the performance of our approach. We start by generating operating characteristics using the weighted risk function in Eq. (18). We use the minimization criterion in Eq. (28) to determine the optimal mean time separating the LGD and LD events, $\mu_X^*$, which we plot versus the mean handover time $\mu_H$. For each value of $\mu_H$, we examined the risk function over a set of $\mu_X$ values that were logarithmically spaced over the range of values from $10^{-4}$ s to $10^6$ s; the value of $\mu_X$ that resulted in a minimum was returned as the value for $\mu_X^*$. We use $\mu_H$ and $\mu_X^*$ to compute operating point values ($P_D^*, P_T^*$), that we plot in a set of operating characteristics. We also develop operating characteristics using the simpler risk function from Eq. 32. Finally, we show how our approach works with a mobile node that moves according to a random walk, which we discussed in Sec. II.

A. Case 1: $C_T \geq C_D$

We first consider the case where $C_T \geq C_D$; in Section III-C, we showed that a solution $\mu_X^*$ exists that satisfies Eq. (28), for all values of the tolerance $\gamma$ and for all values of handover time parameters $a_H, b_H$, and $\lambda_H$ (or $\mu_H$, equivalently). In Fig. 9, we plot $\mu_X^*$ versus $\mu_H$. The figure shows that the sensitivity of $\mu_X^*$ to $C_T/C_D$ increases with $\mu_H$, and that $\mu_X^*$ tends to a finite limit as $\mu_H \to \infty$. The value of the $\mu_X^*$ asymptote decreases as $C_T$ increases with respect to $C_D$, because a larger $C_T$ weight implies that handing over too soon is less desirable than risking a LGD event before the handover completes. Thus, even if the average handover completion time $\mu_H$ is very large, we would resist increasing the threshold for the LGD trigger and thereby increasing $\mu_X$, in order to prevent migrating mobiles from tying up network resources to duplicate and tunnel packets.

To characterize the handover performance for a given value of $\mu_X^*$, we can compute $P_D^*$ and $P_T^*$. Using the values of $\mu_H$ and $\mu_X^*$ from Fig. 9, we show the corresponding operating characteristic plots in Fig. 10, using the same parameter values. For each value of the ratio $C_T/C_D$, we place markers showing the values of $P_D^*$ and $P_T^*$ for the following values of the mean handover time: $\mu_H = 0.1$ s (○), $\mu_H = 0.4$ s (□), and $\mu_H = 6.0$ s (◇). Because reducing $P_T$ is more important than reducing $P_D$, based on our choices of values for $C_T$ and $C_D$, we get curves that are concentrated at the top of the operating characteristic plot. The figure shows that, for the set of parameters that we used, letting $C_T$ increase beyond $2C_D$
produces little change in the operating point $(\mu_H, \mu_X^*)$; we are already near a saturation state. Again we see more variation in the handover performance as $\mu_H$ becomes larger, indicated by the relative spacing of the $\odot$ markers versus that of the $\bigcirc$'s.

**B. Case 2: $\text{C}_D > \text{C}_T$**

Next, we consider the case where $\text{C}_D > \text{C}_T$. We use the same parameters that we used in Section IV-B. We plot $\mu_X^*$ versus $\mu_H$ in Fig. 11 for three values of the ratio $\text{C}_D/\text{C}_T$. We note that the curve associated with $\text{C}_D/\text{C}_T = 1.1$ in the figure is nearly identical to the curve associated with $\text{C}_T/\text{C}_D = 1.1$ that was plotted in Fig. 9, indicating that there is no discontinuity or abrupt change in behavior associated with a change from $\alpha < 1$ to $\alpha > 1$. Note that all the curves in Fig. 11 have vertical asymptotes whose values are given by $\mu_H = \text{C}_T \gamma / (\text{C}_D - \text{C}_T)$, which was predicted by our analysis in Section III-B. Also, we again see the relative insensitivity to the relative sizes of $\text{C}_T$ and $\text{C}_D$ when $\mu_H$ is small.

**C. Results for the Simpler Optimization Criterion**

In this subsection we plot results based on the analysis from Section III-D and compare them to the results from Section IV-A. In all the plots in this subsection, we set $b_H = 0$ s, and we consider two values for $a_H$: 1 and 3. We used the following two values for the tolerance, $\gamma$: 100 ms and 500 ms. Using these parameters, we plot the value of $\mu_X$ that maximizes $\text{Pr}\{H < X \leq H + \gamma\}$ in Eq. (32) in Fig. 13. The figure shows that $\mu_X^*$ is insensitive to the value of $a_H$ when $\mu_H$ is small. Increasing $\gamma$ increases $\mu_X^*$ for a given value of $\mu_H$; interestingly, the sensitivity of $\mu_X^*$ depends strongly on the value of $a_H$. For $a_H = 3$, the sensitivity of $\mu_X^*$ disappears for $\mu_H > 0$, while $\mu_X^*$ remains sensitive to $\gamma$ when $a_H = 1$.

Using the same set of parameter values, we plot $\text{Pr}\{H < X \leq H + \gamma\}$ versus $\mu_H$ in Fig. 14. The probability is close to unity when $\mu_H$ is small and decreases as $\mu_H$ increases. The rate of decrease depends on $a_H$ and $\gamma$; a larger value of $\gamma$, which corresponds to a looser tolerance for the range of times in which $X$ can fall, results in a higher value of $\text{Pr}\{H < X \leq H + \gamma\}$. Also, increasing $a_H$ results in worse performance in the form of a faster rolloff in $\text{Pr}\{H < X \leq H + \gamma\}$. In fact, the rolloff rate increases as $\mu_H$ increases; this effect is more noticeable at larger values of $a_H$. We can see this effect in the figure where the curve associated with $\gamma = 0.5$ s and $a_H = 3$ dips below the curve associated with $\gamma = 0.1$ s and $a_H = 1$ at $\mu_H = 10^4$ s.

We also show an operating characteristic plot in Fig. 15,
we computed \( \mu_H^* \) using Eq. (33), and computed the corresponding minimum risk \( R(\mu_H^*) \) by inserting \( \mu_H^* \) into Eq. (32). We considered two values for the tolerance parameter, \( \gamma \): 100 ms and 2 s. In the simulations, we used the mobility model parameters from Sec. II (these were used to generate Figs. 2–4), which allowed us to use Eq. (11) to compute values of \( R_{LGD} \) from \( \mu_H^* \).

For the simulation runs themselves, we determined the value of \( R_{LGD} \) that minimizes the risk function
\[
1 - \Pr\{H < X \leq H + \gamma\}
\]
We used a range of values from 99.2 m to 99.95 m, in increments of 0.01 m. For each value of \( R_{LGD} \), we performed 500 trials for each of the \( \mu_H \) values that we used; each trial consisted of 50 random walks where we recorded the value of \( X \), the time from the start of the walk at the LDG boundary to the time when the mobile reached the LD boundary. At the end of each random walk, we computed a gamma random variate, \( H \), using the current value of \( \mu_H \), and we recorded where \( X \) fell with respect to the interval \([H, H + \gamma]\) for each walk. At the end of each trial, we estimated \( \Pr\{H < X \leq H + \gamma\} \) using the tallies from the 50 random walks; we used these estimates to get means and standard deviations for the risk function for that pair of \( R_{LGD} \) and \( \mu_H \) values.

In Fig. 16, we show our results for the case where \( \gamma = 0.1 \) s. Fig. 16(a) shows the surface that results from plotting the risk function (i.e. the probability that \( X \) falls outside the range \([H, H + \gamma]\)) versus \( R_{LGD} \) and \( \mu_H \). The shape of the surface reveals easily visible minima with respect to \( R_{LGD} \) for small values of \( \mu_H \); the variation in the risk with respect to \( R_{LGD} \) becomes smaller as \( \mu_H \) increases. Fig. 16(b) shows a contour plot of the surface in Fig. 16(a) using ten contour lines. Superimposed on the contour plot, we use diamonds to plot the values of \( R_{LGD} \) that minimize the risk in the simulations, and we use circles to show the set of points \( \{R_{LGD}(\mu_H^*)\} \) that we obtained from Eq. (33) and Eq. (11).

Comparing the two sets of operating points shown in Fig. 16(b), we observe that our optimization scheme sets the \( R_{LGD} \) boundary an average of 5 cm closer to the \( R_{LD} \) boundary than it would be based on minimizing the risk using the surface shown in Fig. 16(a). For smaller values of \( \mu_H \) (100 ms to 150 ms), this offset is approximately 4 cm. This means that we can expect a higher risk if we use the approach that assumes that \( X \) is exponentially distributed. This is also shown in Fig. 17, which shows the risk that results from using and the mean a minimum risk, versus \( \mu_H \). The figure also shows the value of the risk function in Eq. (32). The figure demonstrates that the difference between the risk associated with computing the LGD boundary and the average minimum risk decreases as \( \mu_H \) increases, such that the risk associated with the computed boundary using \( \mu_H^* \) is consistently within one standard deviation of the optimum risk for \( \mu_H > 300 \) ms. In addition, the risk associated with computing the boundary is with in one standard deviation of the optimized risk function Eq. (32) over the entire range of \( \mu_H \), indicating that the risk function is a reasonably accurate predictor of the true risk.
of performance for this set of parameters.

Next we examine the effect of using a looser tolerance, specifically where $\gamma = 2.0 \text{ s}$, in Fig. 18. We show the risk surface in Fig. 18(a), and the contour plot in Fig. 18(b), along with the values of $R_{\text{LGD}}$ that minimize the risk in the simulations, and the set of points $\{R_{\text{LGD}}(\mu_X)\}$. In this case, we see that the offset between the two sets of $R_{\text{LGD}}$ values is smaller than it was for $\gamma = 0.1 \text{ s}$, and that the offset changes sign as $\mu_H$ increases, with our optimization criterion placing $R_{\text{LGD}}$ about 1 cm farther away when $\mu_H = 100 \text{ ms}$ but placing $R_{\text{LGD}}$ about 1 cm closer for larger values, such as $\mu_H \approx 420 \text{ ms}$.

In Fig. 19, we plot the value of the risk $R(\mu_X)$ from Eq. (32) and the average minimum risk from simulation versus $\mu_H$ for $\gamma = 2.0 \text{ s}$. In this figure, we see less divergence between the optimal risk and the risk that we get from using a $R_{\text{LGD}}$ value developed from $\mu_X$. This agrees with the differences in $R_{\text{LGD}}$ values that were shown in Fig. 18(b). It is interesting to note that both computed risks are significantly less than the theoretical minimum risk, although the divergence is less for smaller values of $\mu_H$. We see from this figure and from Fig. 17 that the risk function in Eq. (32) is a conservative performance estimate and that its deviation from the true risk increases with $\gamma$. We also observe that the deviation between the risk function and the actual risk increases with $\mu_H$ increases when $\gamma$ is large, but that the gap appears to narrow with increasing $\mu_H$ when $\gamma$ is small.

V. SIMULATION RESULTS

In this section, we show results from a series of simulations to illustrate how one can experimentally compute values for the LGD threshold, $T_{\text{LGD}}$, based on the risk functions described in Section III. We used the ns-2 tool [10] to do packet-level simulations of a handover involving a single mobile node migrating from one AP coverage area to another. We show the topology of the simulated network in Fig. 20. Because of the short distances used in this scenario, we use the free space path loss in the link budget computations, which is

$$L_{\text{path,db}}(\ell) = 20 \log_{10} \left( \frac{4\pi \ell f}{c} \right), \quad (34)$$

where $\ell$ is the path length in meters, $f = 2.412 \text{ GHz}$ is the carrier center frequency, and $c = 2.997 \times 10^8 \text{ m/s}$ is the propagation speed of the carrier wave in air. We do not consider fading or shadowing in these simulations.
Fig. 18. (a) Mesh plot of risk with $\gamma = 2.0$ s, with each point computed by averaging the results from 500 runs of 50 random walks each, plotted versus mean handover time, $\mu_H$, and LGD boundary distance, $R_{\text{LGD}}$. (b) Contour plot of mean risk versus $\mu_H$ and $R_{\text{LGD}}$, with overlay plots showing $R_{\text{LGD}}^*(\mu_X^*)$ (circles) and $R_{\text{LGD}}$ values that correspond to minima on the risk surface for a given value of $\mu_H$ (diamonds).

Fig. 19. Plot of theoretical risk metric and measured risk, using computed and optimal $R_{\text{LGD}}$ values, versus mean handover time, $\mu_H$, for $\gamma = 2.0$ s.

The mobile and the APs communicate using IP in the network layer over IEEE 802.11 at the link layer. The transmitter power of each AP is 100 mW. We use isotropic transmit and receive antennas (unity gains), and assume no system losses. The radius of each AP’s coverage area is 50.0 m; a mobile at that distance would experience a RSS of $-84.069$ dBW, from Eq. (34). When the mobile’s RSS falls to this level, it causes a LD trigger event and the connection with AP1 breaks. The central AP, AP1, is surrounded by six other APs as shown in Fig. 20. The coverage limit of AP1 is shown in red in the figure. All seven APs are connected to a single router, which is also connected to the mobile’s MIPv6 Home Agent (HA) and the Corresponding Node (CN), with which the mobile is communicating. The data rate on all of the wired links from the router to other entities in the network is 100 Mb/s.

We examined a range of values of $R_{\text{LGD}}$ from 40.0 m to 47.75 m, at intervals of 0.25 m. Each value of $R_{\text{LGD}}$ corresponds to a unique value for the LGD threshold, $T_{\text{LGD}}$, where, from Eq. (34),

$$T_{\text{LGD}} = -10 \text{ dBW} - L_{\text{path}, \text{dB}}(R_{\text{LGD}}).$$

We performed 18 000 runs for each value of $T_{\text{LGD}}$ and recorded the results in an ASCII output file; for each run we recorded the time when the LGD trigger fires, the time when the LD trigger fires, and the time when the handover completes. Each simulation run begins at time $t = 0.0$ s with the mobile located close to AP1 so that it can perform all the signaling to attach itself to the network. The mobile then moves outward from the center of API’s coverage area so that at $t = 5.0$ s, the mobile

Fig. 20. Network topology used in ns-2 simulations.
is located on the circle with radius \( R_{LGD} \) centered on AP1, with a phase angle that is uniformly distributed over the interval \([0, 2\pi]\). The mobile’s initial velocity vector is aligned with the radius vector from AP1 to the mobile, i.e. it is moving away from AP1.

In each simulation run, the mobile’s LGD trigger fires once the mobile detects a beacon signal that allows it to determine that its RSS has fallen to \( T_{LGD} \); the mobile immediately begins scanning for a target network at the link layer, and uses RSS from each of the six other APs to decide which one will be its target AP. We record the LGD time, \( T_{LGD} \), in the output file. Because this simulation uses the free-space path loss model, the chosen AP will be the one whose distance to the mobile is the smallest. Once the mobile has selected a target AP, it uses FMIPv6 predictive signaling to set up a tunnel between the target access point and AP1, followed by MIPv6 signaling between the mobile and the HA. The handover time is random, with mean packet delays of 50.0 ms and 30.0 ms between the APs and the router and between the router and the HA, respectively. The are no packet losses on the wired links. When the handover completes, the time \( t_{handover} \) is recorded and stored in the output file.

The mobile moves according to a random walk model while maintaining a constant speed of 2.0 m/s. Every 0.5 s, the mobile changes direction by picking an angle that is uniformly distributed over the interval \([-\pi/20, \pi/20]\). It adds this to the current angle of its velocity vector and then moves 1.0 m in the new direction. The mobile repeats this process until its distance from AP1 reaches 50.0 m or until the run time reaches 200.0 s; either criterion causes the run to end. If the run ends because a LD event occurs, we record the time, \( t_{LD} \), in the output file. If the run ends because the time limit was reached without the mobile’s RSS falling to the LD threshold, the LD time for the run is recorded as \(-1\).

For each value of \( T_{LGD} \), we obtained sample values \( \hat{h} \) and \( \hat{x} \) of \( H \) and \( X \), respectively, for each of the 18,000 runs that did not end in a timeout by computing \( \hat{h} = t_{handover} - t_{LD} \) and \( \hat{x} = t_{LD} - t_{LGD} \). We calculated an estimate of the risk function for a given value of \( \gamma \) by computing the relative frequency of the probabilities \( P_D = N_{x\leq h}/N_{LD} \) and \( P_T = N_{x<h+\gamma}/N_{LD} \), where \( N_{x\leq h} \) is the number of runs in which the recorded handover duration was less than the recorded time between the LGD and LD events, \( N_{x<h+\gamma} \) is the number of runs in which the recorded handover duration was less than the sum of \( \gamma \) and the recorded time between the LGD and LD events, and \( N_{LD} \) is the number of runs that ended in a LD event. In all the simulations that we performed, the smallest value of \( N_{LD} \) was 17,125 when we used \( R_{LGD} = 40.25 \) m; this value corresponds to a 95.1% probability of a run terminating in a LD event rather than a timeout. The resulting estimated risk is

\[
\hat{R}(T_{LGD}, \gamma) = C_D \hat{P}_D + C_T (1 - \hat{P}_T),
\]

and the estimated variance of the sample risk [11] is

\[
s^2_R = \frac{\hat{R}(T_{LGD}, \gamma)(1 - \hat{R}(T_{LGD}, \gamma))}{N_{LD} - 1}.
\]

In Fig. 21, we plot the estimated risk function versus \( T_{LGD} \) for four values of \( \gamma \), with 99% confidence intervals shown.

In Fig. 22, we plot the simple risk function defined in Eq. (31), which is Eq. (18) with \( C_D = C_T \). We show the risk for the full range of LGD threshold values that we considered, and for various values of \( \gamma \) on a logarithmic scale from \( 10^{-2} \) s to \( 10^1 \) s. If we take the minimum risk associated with each value of \( \gamma \), we find that all of them occur at the same value of \( T_{LGD} = -83.485 \) dBW. This threshold corresponds to a distance of 46.75 m from AP1, or 3.25 m inward from the maximum coverage radius.

Because the optimal LGD threshold is insensitive to \( \gamma \) when the weights for \( P_D \) and \( 1 - P_T \) are equal, we want to examine the sensitivity to \( \gamma \) when the weights are not equal. We varied the ratio \( C_T/C_D \) from \( 10^{-2} \) s to \( 10^2 \) s and computed the value of \( T_{LGD} \) that minimized the weighted risk for each value of \( \gamma \) from \( 10^{-2} \) s to \( 10^1 \) s. We plot the resulting optimal values of \( T_{LGD} \) in Fig. 23. The figure confirms that the value of \( T_{LGD} \) that minimizes the risk function is constant with respect to \( \gamma \) when \( C_T = C_D \). The figure also shows that the optimal threshold is insensitive to \( \gamma \) for a broad range of values of \( C_T/C_D \), which we expect since \( P_T \) is deemphasized when \( C_T/C_D \) is very small and \( P_D \), which does not depend on \( \gamma \), becomes the dominant component of the risk function.

In both the surface and contour plots, a steep transition
in the optimal value of $T_{\text{LGD}}$ is clearly visible; this transition is associated with a critical value of the weight ratio $C_T/C_D$ for a given value of $\gamma$. When $\gamma$ is very small, this critical value of the weight ratio is around four or five as shown in Fig. 23(b); as $\gamma$ increases, the critical value of the weight ratio increases to around 10. For values of $C_T/C_D$ above the critical value, the optimal value of $T_{\text{LGD}}$ is $-83.669$ dBW when $\gamma < 1.0$ s. When $C_T/C_D$ is greater than the critical value and $\gamma$ is larger than 1.0 s, the optimal value of $T_{\text{LGD}}$ is $-83.623$ dBW, as shown in Fig. 23(a); this occurs when the ordered pair $(C_T/C_D, \gamma)$ lies within the region in the upper right portion of Fig. 23(b). This larger threshold value is associated with the relaxation of the handover promptness criterion measured by $P_T$ that occurs when $\gamma$ is large. An interesting result from this figure is that over a large portion of the parameter space, we would use one of two possible values of $T_{\text{LGD}}$, depending on the relative importance of the performance criteria associated with $P_D$ and $P_T$.

VI. SUMMARY

In this paper, we developed a weighted risk function to characterize the cost associated with using a particular value of $\mu_X$, the mean time between the LGD and LD trigger events. We also developed a simplified risk function that expressed the risk is the probability that the Link Down event would occur outside the time range $[H, H+\gamma]$, and that is a special case of the weighted risk function, where the weights are equal. Both risk functions assume that $X$ has an exponential distribution. We showed that for the case where the mobile’s path can be modeled as a random walk and the LGD boundary is close to the LD boundary, this assumption is reasonable.

For our first set of experimental results, we used a shifted gamma distribution to model the handover time and showed how the resulting $\mu_X$ value varies with respect to the mean handover time, $\mu_H$. We created operating characteristics to examine the effect of the value of $\mu_H$ and the risk function weights $C_D$ and $C_T$ on handover performance. We also simulated the performance of a handover scheme that used our simplified risk function to obtain a value for the distance from the AP to the boundary where the Link Going Down event is triggered. We showed that our approach provides results that are nearly optimal even when $X$ is not exponentially distributed, although the deviation from the optimal performance depends on various parameter values such as the tolerance variable $\gamma$.

Finally, we used a packet-level simulation of a mobile migrating from one access point to another to show how the risk function can be used to determine the optimal LGD threshold for a given set of cost weights $C_T$ and $C_D$ and a given tolerance $\gamma$. We showed that for the example that we considered, the optimal threshold value is insensitive to changes in the weights and tolerance over large regions of the parameter space. Using this approach, network operators can devise thresholds for their event triggers that optimize handover performance based on the relative importance of prompt handovers versus LGD trigger sensitivity.

ACKNOWLEDGMENTS

The authors thank Richard Rouil of NIST for running the ns-2 simulations for Section V.

REFERENCES

Fig. 23. (a) Mesh plot of optimal value of $T_{\text{LGD}}$ versus $\gamma$ and $C_T/C_D$. (b) Contour plot of optimal value of $T_{\text{LGD}}$ versus $\gamma$ and $C_T/C_D$, showing boundaries between regions where $T_{\text{LGD}}$ is constant.