Adaptive Predictive Distance-Based Mobility Management for PCS Networks[1]

William Navidi and Tracy Camp

Department of Mathematical and Computer Sciences
Colorado School of Mines
Golden, Colorado 80401

Abstract

Mobility management methods that predict the location of a mobile from accurate knowledge of its previous position and its velocity can reduce costs considerably over non-predictive methods that always begin paging in the last known location. When the mobility pattern is well-described by a tractable mathematical model, such as the Gauss-Markov model, a predictive management scheme based on that model is ideal. However, in many cases, mobility patterns are too complex to be modeled accurately. For use in these situations, we describe an adaptive method, in which the location of a mobile is predicted solely on the basis of previously reported locations, *without specifying a mathematical model for mobility*. Simulations show that our adaptive method performs well under a wide variety of mobility patterns, while a method that predicts location under the assumption of a Gauss-Markov mobility model can perform poorly when the model assumptions are violated. While our method is simple, we show that it (almost) always performs better than a previously published method that is more complex. In addition, our adaptive method outperforms a non-predictive method in almost all situations as well.

Keywords: personal communications systems, mobility management, location prediction, adaptive methods

1 Introduction

As technology has matured, computers have become more and more ubiquitous in our society and we have seen tremendous growth in the popularity of laptop computers and personal digital assistants (PDAs). As portable devices, these computers are often carried and used in cars, in coffee houses, and in other non-traditional places. For example, consider the number of laptop computers used on airplanes today compared with the number that were used a decade ago. The desire to stay connected now goes well beyond carrying a pager and a cellular telephone.

[1]This work was partially supported by National Science Foundation grant ANI-0208352. Research group's URL is *http://toilers.mines.edu*. Reference for this manuscript: Technical Report MCS-03-07, The Colorado School of Mines, May 2003.

It is, therefore, no surprise that the users of these portable devices desire Internet access anytime and anywhere. While wireless support is available for portable computers, it is still very sparsely installed. And, of course, mobility poses additional problems for computer communications. Forman and Zahoran have classified these challenges as follows [1]:

- Wireless issues: compared to traditional wired links, wireless communication has lower bandwidths, lower reliability (or higher error rates), frequent and possibly sporadic disconnections and blackouts, as well as problematic security.
- Portability issues: compared to traditional static computers, portables have lower power capacity.
- Mobility issues: tracking mobile users that may move stochastically.

Our work has been dealing with the last challenge above, i.e., the problem of online tracking of mobile users (or mobiles), known widely as the location management problem. Location management is an important problem in both cellular telephone systems and mobile computing and networking systems.

A personal communications service (PCS) network consists of mobile users moving within a specified coverage area. In order to provide efficient use of the radio spectrum, a cellular architecture is employed. Each cell is served by a base station. Mobile users rely on the Base in its current cell for connectivity to the network. For example, to transmit a message, the mobile sends the message to its current Base; the Base then routes the message on the wired network to the destination (static host or Base unit where the destination is currently located) of the message. Figure 1 illustrates a PCS network.

When a call arrives, the network must locate the cell in which the receiving mobile unit is currently located; this process is known as terminal paging. In order to reduce paging costs, mobiles inform the network of their locations from time to time; this is known as location updating. There is a tradeoff between location update costs and paging costs. If a mobile expends power and bandwidth to update its location more often, it can reduce the area that needs to be paged when a call arrives. On the other hand, if updates are performed less frequently, more power and bandwidth will be expended on paging, since larger areas will need to be paged.

The cost of mobility management over any given time period is the sum of the cost of the location updates and the cost of paging for the calls that arrive during that time period. Much research has

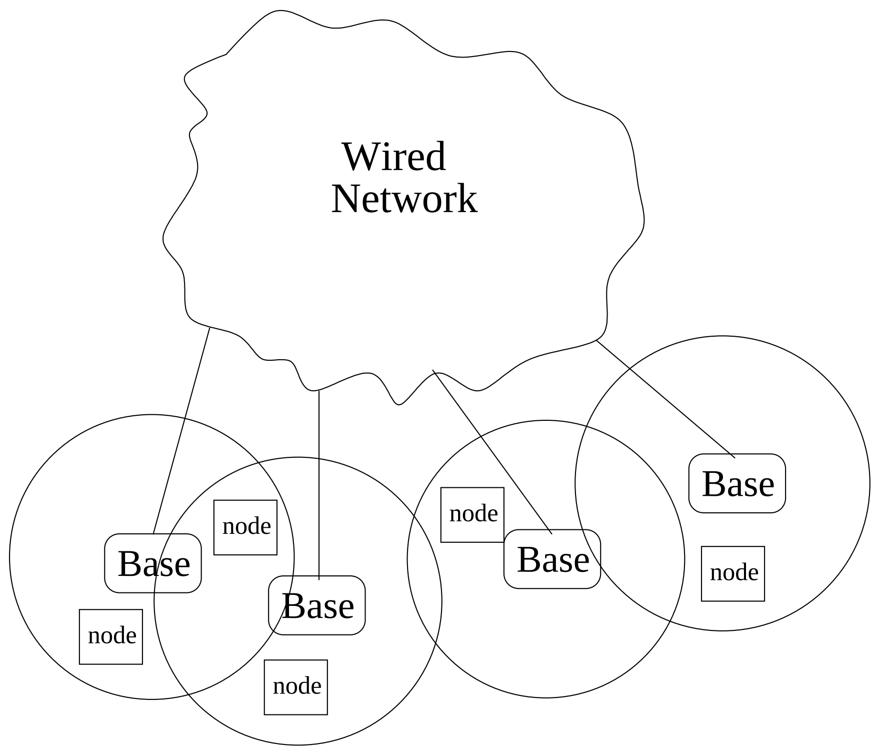

Figure 1: An example PCS network.

therefore focused on the development of algorithms to reduce the sum of these two costs. In *static* mobility management algorithms, the coverage area is divided into several location areas, each of which contains a group of cells. A mobile updates its location whenever it crosses a location area boundary. When a call arrives, the network simultaneously pages all the cells in the location area where the mobile is located.

In contrast, *dynamic* mobility management algorithms [5, 6] in general do not involve location areas, and paging is done one cell at a time. Dynamic mobility management algorithms therefore require both a location update algorithm, to determine when updates will occur, and a paging algorithm, to determine the order in which cells will be paged. If, when an incoming call arrives, the network can determine a probability distribution for the cell in which the mobile is located, the optimal paging scheme is to page in order of decreasing probability.

In section 2, we discuss a variety of location update methods that appear in the literature. We then give details for one predictive method that assumes the velocity of the mobile follows the Gauss-Markov model. Finally, in section 2, we propose a new simple location update method which we illustrate performs (almost always) better than a more complex method. In section 3, we compare the performance of our model to the Gaussian method, and to the non-predictive method, in a situation where the Gauss-Markov mobility model holds. In section 4 we make the same comparisons in a situation where the Gauss-Markov model does not hold. In section 5, we make the comparisons in a situation where the mobility follows a Gauss-Markov process for a random, geometrically distributed time, then initiates a new, independent process by choosing a velocity independent of the previous one. In other words, a geometric renewal process is superimposed on the Gauss-Markov process. This is meant to model a mobile that reaches a destination from time to time, then moves away at a velocity independent of the velocity with which it reached the destination. In section 6, we make the comparisons in a setting where the Gauss-Markov model holds, but the network misspecifies the parameters of the model. In section 7 we summarize our results and state our conclusions.

2 Mobility management methods

2.1 Related work

A variety of algorithms have been proposed for location updates. These include *timer-based* schemes, *movement-based* schemes, and *distance-based* schemes. In timer-based schemes [4, 5], updates are made at fixed time intervals. In movement-based schemes [3, 4], updates are made whenever the number of cell-boundary crossings since the last update exceeds a specified threshold. In distance-based schemes [4, 7], a mobile updates its location whenever its distance from an expected location exceeds a specified value. This paper considers distance-based methods.

In the earliest distance-based methods, the “expected” location is simply the location at the last update. In other words, a mobile updates its location whenever its distance from its last updated location exceeds a specified value, and paging always starts at the last updated location. Such methods can be called *non-predictive*, because no prediction is made of the distance traveled since the last update. In contrast, *predictive* mobility schemes [2, 10, 11] require some knowledge of mobility patterns, based either on prior knowledge or on movement history.

A *predictive* distance-based mobility management scheme that is based on prior knowledge of a stochastic model is proposed in [2]. In this method, the mobile terminal reports both its location and its velocity when updating, At fixed time intervals, the network uses a stochastic model for the velocity of a mobile to compute a probability density function for the mobile’s location. Upon call arrival, the network pages the mobile, starting from the most probable location and continuing in decreasing order of probability. For many velocity models, the density is symmetric around its expected value. In these cases, the optimal paging strategy is the shortest-distance-first strategy, in which the network first pages the expected location, then moves outward, always paging the cell nearest the expected location among those not yet paged. In [10], mobiles construct a user mobility profile (UMP), which is a sequence of pairs consisting of cell identification and expected entry times. The UMP is derived from a user mobility history (UMH), which is a data structure where the mobile stores its mobility history. When a call arrives, cells are paged sequentially, starting from the cell whose expected entry time is the most recent time prior to the current time. In [11], the idea of a prediction area is discussed. Each mobile has its own individually constructed prediction area, determined by its mobility pattern. The prediction area defines the limits of paging for a mobile, in other words, the prediction area consists of the largest number of areas that will be paged. In [12], a method of location management appropriate for networks with mobile bases is presented. A robust extended Kalman filter is used to predict the arrival time of a mobile in the next cell. Measurements from only two base stations, fixed with respect to the current cell, are needed. Location and acceleration of mobile bases are assumed known via GPS, while no such information needs to be available about a mobile. In [13], the use of a genetic algorithm is described, to determine a selective update strategy that is tailored to the specific characteristics of the network. In [14], a scheme is described in which the decision to update location is based on both a distance and a time threshold. Both of these thresholds can be adapted to the changing velocity of the mobile. Expected update and paging costs are calculated on the assumption that call arrivals follow a Poisson process.

2.2 The Gaussian method

In [2], the predictive distance-based mobility management scheme was studied under the assumption that the velocity of a mobile followed a one dimensional discrete-time Gauss-Markov model. Because of this assumption, we will refer to this predictive method as the Gaussian predictive method, or more simply, as the *Gaussian method*. To describe the Gauss-Markov model, let $v(t)$ denote the velocity of the mobile at time t, and let μ denote the asymptotic mean velocity. The Gauss-Markov model is stationary, so the correlation between $v(t-1)$ and $v(t)$ is the same for all t. Let α denote this correlation. Let $\varepsilon(t)$ denote an independent and identically distributed Gaussian process with mean 0 and variance σ^2. The discrete-time Gauss-Markov model is given by

$$v(t) = \alpha v(t-1) + (1-\alpha)\mu + \sqrt{1-\alpha^2}\,\varepsilon(t) \tag{1}$$

At each update, we will set $t = 0$, so that $v(0)$ represents the velocity at the last update, and we will reset the origin to the location of the mobile. Then the expected location of the mobile t time units after the last update is given by

$$l_{exp} = t\mu + \frac{1-\alpha^t}{1-\alpha}(v(0)-\mu) \tag{2}$$

In the Gaussian method, paging starts at l_{exp} and proceeds in a shortest-distance-first fashion. If the movement of the mobile does indeed follow the Gauss-Markov model (Equation 1), then this is the optimal paging strategy. To implement this strategy, the network must know the values of the parameters α and μ, and must receive both the location and the velocity of the mobile at each update. In [2], it is shown that when the Gauss-Markov model holds, this strategy in general far outperforms a non-predictive distance-based strategy.

The superiority of the Gaussian method is due to network knowledge of the mobility model, and in particular to knowledge of the quantities μ, α, and $v(0)$. When the velocity of the mobile is truly governed by the Gauss-Markov model (Equation 1), a non-predictive distance-based method can do no better than the Gaussian predictive method. Of course, when mobility patterns are too complex to be described by a stochastic model, the Gaussian predictive model is not directly applicable. It turns out, however, that much of the benefit of the predictive method can be attained without specifying a model for mobility. Table 1 summarizes our notation.

2.3 Our adaptive method

We propose an adaptive predictive distance-based method, which we refer to as the *adaptive method*. In this method, no mobility model is assumed for the network. At each update, the mobile transmits only its location to the network. The network then computes the distance traveled and the time elapsed since the last update, and determines the average velocity during that time. When an incoming call arrives, the network computes a predicted position for the mobile, based on the assumption that the mobile has been moving at the most recently computed average velocity. Therefore, if l_0 is the location at the last update, and $\bar{v}$ is the average velocity in the time interval between the last two updates, the predicted location t time units after the last update is

$$l_{exp} = l_0 + \bar{v}t \tag{3}$$

Table 1: Summary of Notation

$v(t)$	velocity at time t
α	correlation between $v(t-1)$ and $v(t)$ in Gauss-Markov model
$\varepsilon(t)$	random error in $v(t)$ in Gauss-Markov model
μ	theoretical mean velocity in Gauss-Markov model
σ	standard deviation of $v(t)$ in Gauss-Markov model
l_{exp}	expected location of mobile
$\bar{v}$	average velocity observed since last update
C_u	cost of update when cost of page is set to 1
m	location inspection interval
λ	call arrival rate
λ_{renew}	rate at which new excursions begin in renewal model
τ	standard deviation of mean velocity for an excursion in renewal model
μ_i	theoretical mean velocity for the ith excursion in renewal model
α^*	value of correlation used for updates in misspecified model
μ^*	value of mean velocity used for updates in misspecified model

The network pages the mobile first at location l_{exp}, then proceeds in a shortest-distance-first fashion. Location updates are performed according to the algorithm in [2]. The mobile checks its location periodically, and updates its location whenever its distance from the predicted location exceeds a specified threshold.

Our method is a hybrid between a non-predictive distance-based method and the Gaussian method. Like the non-predictive method, mobiles update only their locations, and no mobility model is assumed by the network. Like the Gaussian method, the network begins paging from a predicted location that is computed when a call arrives, rather than from the most recently updated location.

The adaptive method is designed to outperform the non-predictive distance-based method in situations where velocities are correlated in time, without requiring that they follow a specific stochastic model. In addition, in situations where mobility is actually governed by a known stochastic process, the adaptive method should perform nearly as well as a method which assumes knowledge of that process.

We note that an adaptive algorithm has been proposed [8] for choosing a distance threshold for location updates in the context of the non-predictive distance based method. Our method does not choose a distance threshold; instead it provides a method for predicting location without assuming a mathematical model for mobility. Bhattacharya and Das [9] developed a method that assumes that mobility follows a stationary stochastic process, but does not require a Gaussian assumption.

Finally, we note that this model can be applied to movement in one, two, or three dimensions, since l_{exp}, l_0, and $\bar{v}$ can be vectors as well as scalars.

2.4 Implementation of the methods

We assume that the network consists of a one-dimensional array of cells extending infinitely in each direction. The unit of distance is taken to be the width of a cell. We assume that a location inspection time period of length m is defined. After each m time units, a mobile determines its current location and computes its expected location. For the non-predictive method the expected location is the location at the last update. For the Gaussian method, the expected location is given by Equation 2. For the adaptive method, the expected location is given by Equation 3. If the distance between the current location and the expected location exceeds a specified threshold D, an update is performed. In the Gaussian method, both location and velocity are updated. In the other two methods, only location is updated.

We assume that calls arrive according to a geometric random process, that is, that the waiting time between calls are independent, identically distributed geometric random variables. We denote the rate at which calls arrive by λ, so that the expected waiting time between calls is $1/\lambda$. For simplicity, we assume, as in [2], that all calls arrive at location inspection times. If the criteria for an update are met at the time that a call arrives, we assume that the update is performed just before the call arrives. We define the cost of paging a cell to be 1. We denote the cost of an update by C_u. We assume that the cost of updating both location and velocity is the same as the cost of updating location alone. Finally, whenever a call arrives, the location of the mobile is determined by the paging process. Therefore a location update automatically occurs immediately after every call, whether or not the expected location exceeds the distance threshold D.

3 Comparison when the Gauss-Markov model holds

We measured the total cost of the non-predictive method, the Gaussian method, and the adaptive method through simulation. In each simulation, we generated a realization of the Gauss-Markov model over a time period that spanned 10,000 calls. We performed simulations for various values of six parameters. Three of them, α, σ, and μ, specify the mobility model (Equation 1). We also varied the location inspection interval m, the call arrival rate λ, and the cost per update C_u (the cost of paging a cell is set at 1). For each set of parameters and each of the methods, the distance threshold D was chosen to minimize the total cost (see below). In other words, we compared each method at the value of D that provided its best performance. Following [2], we define the performance gain of the Gaussian or the adaptive method to be the quotient of the total cost of the non-predictive method divided by the total cost of the Gaussian or adaptive method, respectively.

The range of values for D is shown in Table 2. For most of the parameter choices, the values were between 2 and 5 cells for all three methods. The few large values occurred when update costs were high or when the call rate was very low. The reason that a large threshold is optimal when the call rate is low is that when calls seldom come in, pages are few, so page costs will be low even when few updates are done.

Table 2: Values of D

Method	Lowest value	Highest value
Gaussian	1	16
Adaptive	1	15
Non-predictive	1	36

The performance gain of the Gaussian method compared to that of the adaptive method did not vary much with m (results not shown). Figures 2–6 present performance gains of the Gaussian and of the adaptive methods for a variety of choices of the remaining parameters. In general, the performance of the adaptive method is close to that of the Gaussian method, and in most cases much better than that of the non-predictive method. Figure 2 plots performance gain vs. α, where α takes the values 0, 0.5, 0.75, 0.9, and 0.99. It can be seen that the advantage of either method over the non-predictive method is non-monotonic in α. The reason for this, given in [2], is that the variance of the distance traveled since the last update (denoted C_{kk} in [2]), is non-monotonic in α, and in our simulations, (as well as in those reported in [2]), reaches a maximum for a value of α

near 0.9. The shape of the curve in Figure 2 reflects the fact that the gain due to prediction is a decreasing function of this variance. In principle, the advantage of the Gaussian method over the adaptive method should decrease with this variance as well. A slight decrease can be observed in Figure 2.

Figure 3 plots performance gain vs. σ, where σ takes the values 0.1, 0.5, 1.0, and 5.0. The advantage of the Gaussian method over the adaptive method, and the gain of either method over the non-predictive method, is a decreasing function of the error standard deviation σ. This is the same phenomenon observed in Figure 2, although somewhat more pronounced. When the error standard deviation is equal to the mean velocity ($\sigma = 1$), the Gaussian method is scarcely better than the adaptive method. When $\sigma = 5$, the Gaussian method offers essentially no advantage, and neither predictive method offers much advantage over the non-predictive method.

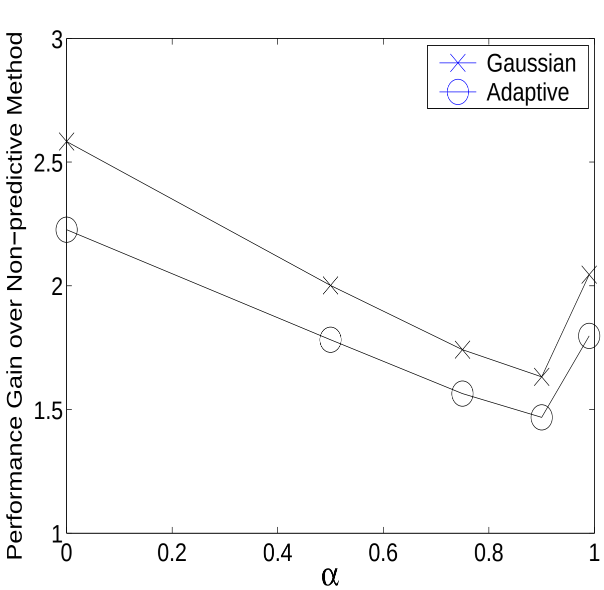

Figure 2: Performance gain as a function of the correlation α. Values of the other parameters are $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.007 and 0.022. All differences between Gaussian and adaptive are statistically significant at the 1% level.

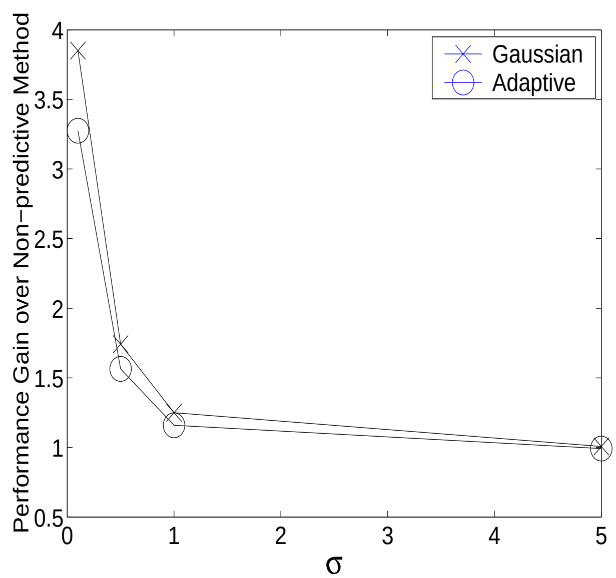

Figure 3: Performance gain as a function of the standard deviation σ. Values of the other parameters are $\alpha = 0.75$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.002 and 0.027. All differences between Gaussian and adaptive are statistically significant at the 1% level.

Figure 4 plots performance gain vs. μ, where μ takes the values 0, 0.5, 1.0, and 5.0. The advantage of the Gaussian method over the adaptive method increases somewhat with the mean velocity

μ. This is due to the fact that μ is known in the Gaussian method and estimated in the adaptive method. The advantage of precise knowledge of the mean velocity is greater when the mean velocity is greater. Note that when $\mu = 0$, the non-predictive method outperforms the adaptive method, and is nearly as good as the Gaussian method. The reason for this is that the so-called non-predictive method actually predicts a mean velocity of 0, since it assumes that the location has not changed since the last update. When the mean velocity actually is 0, the non-predictive method therefore does quite well. The slight advantage for the Gaussian method in this situation is due to its knowledge of the correlation α. To put the units in Figure 4 in perspective, note that the location update period is taken to be $m = 10$. Therefore, when $\mu = 5$, the average distance traveled between location inspections is 50 cells.

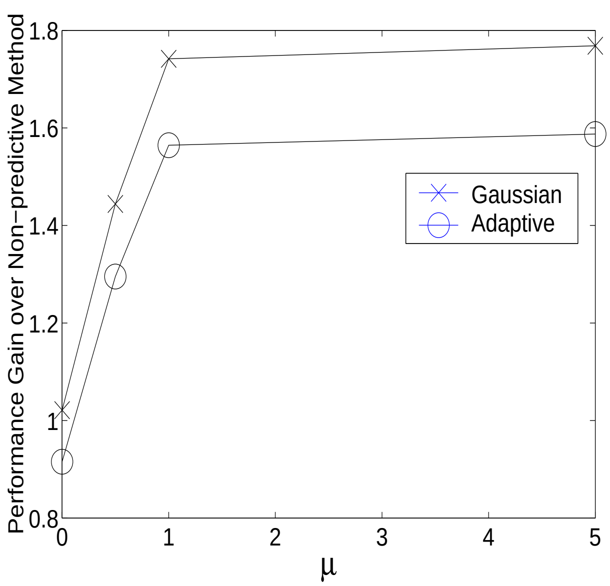

Figure 4: Performance gain as a function of the mean velocity μ. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.004 and 0.015. All differences between Gaussian and adaptive are statistically significant at the 1% level.

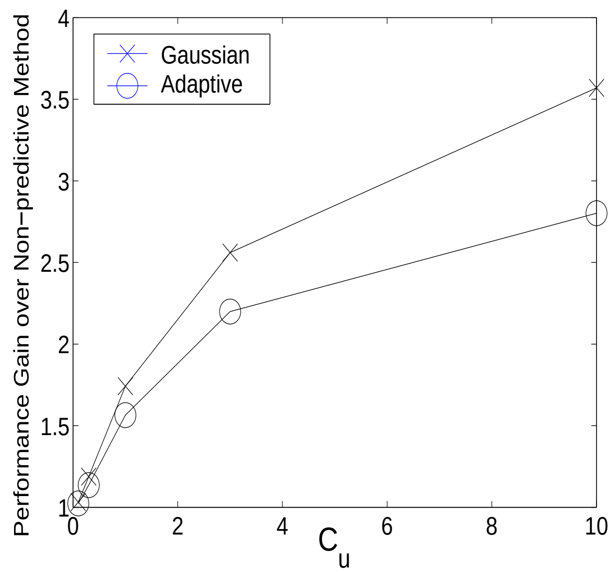

Figure 5: Performance gain as a function of the update cost C_u. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.003 and 0.028. The difference between Gaussian and adaptive at $C_u = 0.1$ is not statistically significant at the 5% level; all other differences are statistically significant at the 1% level.

Figure 5 plots performance gain vs. the update cost C_u, for values of C_u equal to 0.1, 0.3, 1.0, 3.0, and 10.0. When the update cost is very small, the optimal strategy is to update at every opportunity. When updates are frequent, the advantage of prediction is less, so the predictive methods are essentially no better than the non-predictive method. As C_u increases, so does the advantage of prediction, since the predictive methods require fewer updates. The advantage of the Gaussian method over the adaptive method increases with C_u as well, for the same reason.

Figure 6 plots performance gain vs. the call frequency λ for values of λ equal to 0.05, 0.01, 0.005, and 0.001. The advantage of prediction decreases as the call frequency increases. The reason for this is that when a call arrives, the location of the mobile is determined, which results in a "free" update. Thus, when calls arrive frequently, the update cost is effectively reduced. This lessens the advantage of predictive methods over the non-predictive method, and of the Gaussian method over the adaptive method, for the reasons explained in the discussion of Figure 5.

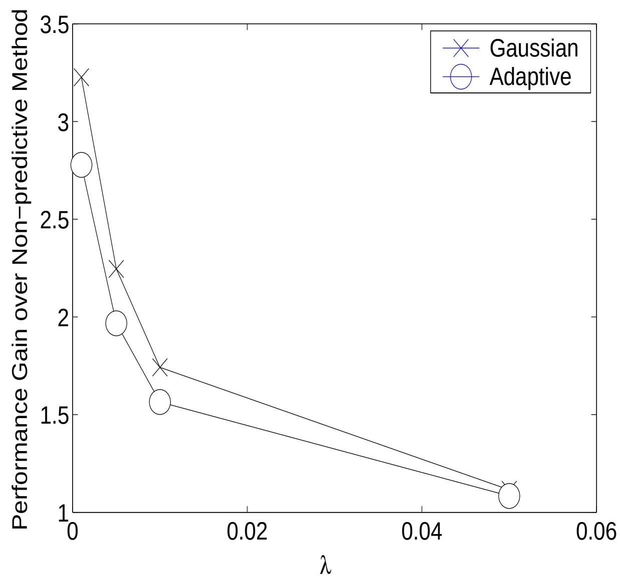

Figure 6: Performance gain as a function of the call frequency λ. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, and $C_u = 1.0$. Standard errors for the performance gains range between 0.003 and 0.014. All differences between Gaussian and adaptive are statistically significant at the 1% level.

4 Comparison when the Gauss-Markov model does not hold

An essential feature of the Gauss-Markov model is that the mean velocity of a mobile is constant over time. Since the Gaussian method assumes that mobility follows the Gauss-Markov model, it does not perform well when the assumption of constant mean velocity is seriously violated.

We compared the performances of the Gaussian, adaptive, and non-predictive methods through simulation, in the context of a mobility model where the speed of the mobile begins at 0, increases for a time, then decreases at the same rate back to 0. This is designed to model a mobile that travels to a destination by systematically accelerating over the first half of the trip, then systematically decelerating over the second half. As in section 3, we generated, for each set of parameters, a realization of the process described above that spanned 10,000 calls. The parameters governing the performance are the same six that were discussed in section 3.

The mobility model used in these simulations is the sum of a Gauss-Markov process (Equation 1) with $\mu = 0$ and a deterministic process γ defined below. In the definition, N is the total length of the simulation.

$$\gamma(t) = \begin{cases} 20t/N & t \leq N/2 \\ 20 - 20t/N & t > N/2 \end{cases} \tag{4}$$

The value of $\gamma(t)$ is 0 at $t = 0$. It then increases linearly to 10 at $t = N/2$, after which it decreases linearly to 0 at $t = N$. In our simulations, since the maximum speed of the mobile was set to 10, the mean speed over the length of the simulation was 5. To implement the Gaussian method, we used Equation (2) with $\mu = 5$, which is the true mean velocity over the distance traveled by the mobile in our simulation. We used $\mu = 5$ in all our simulations, since the true mean should be the best value for the Gaussian method.

Figures 7–10 present the simulation results. The performance of the Gaussian method under this mobility model is much worse than it is under the Gauss-Markov model; in fact it is negligibly better than the non-predictive method unless the correlation is near 1. In contrast, the performance of the adaptive method under this mobility model is as good or better, relative to the non-predictive method, than it is under the Gauss-Markov model (compare Figure 7 with 2, 8 with 3, 9 with 5, and 10 with 6).

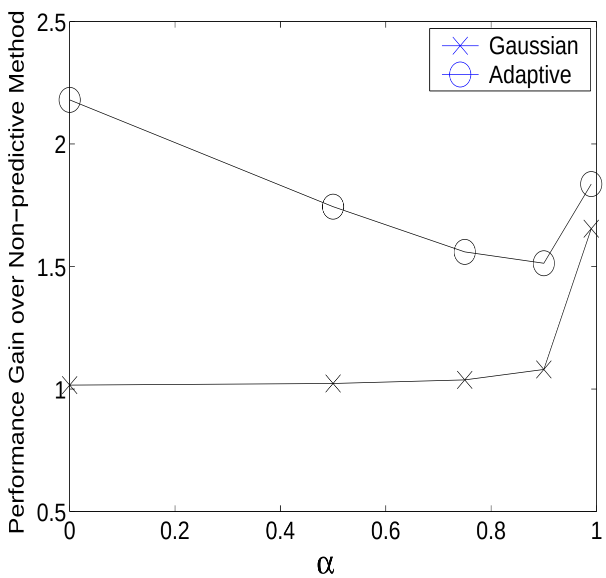

Figure 7: Performance gain as a function of the correlation α. Values of the other parameters are $\sigma = 0.5$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.001 and 0.016. All differences between Gaussian and adaptive are statistically significant at the 1% level.

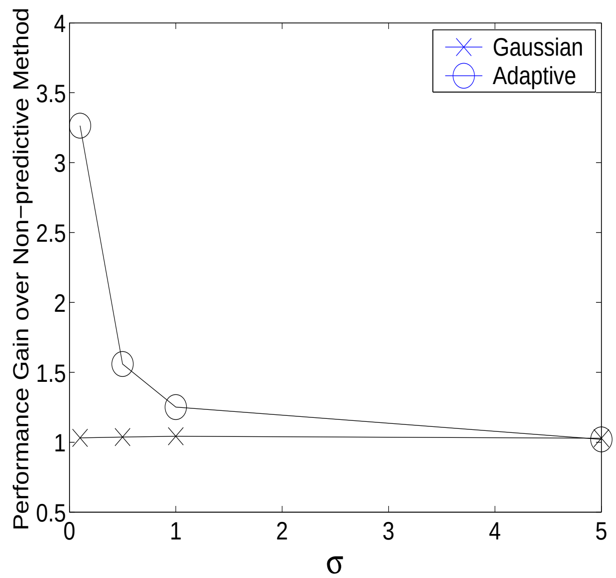

Figure 8: Performance gain as a function of the standard deviation σ. Values of the other parameters are $\alpha = 0.75$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.0003 and 0.018. All differences between Gaussian and adaptive are statistically significant at the 1% level.

5 Comparison under a Gauss-Markov model with renewal

We now compare the methods in the context of a mobility model that is designed to mimic the behavior of a mobile terminal that reaches a sequence of destinations as it travels. The portion of the travel that occurs between two consecutive destinations will be referred to as an *excursion*. The movement of the mobile therefore consists of a sequence of excursions. In our model, the durations of the excursions are independent geometric random variables. Within any excursion the velocity of the mobile follows the Gauss-Markov model. The mean velocity varies from excursion to excursion. Specifically, for each excursion, the mean velocity is chosen independently from a Gaussian distribution.

To describe the model precisely, let $t_1, \ldots$ be the times at which destinations are reached. The differences $t_1 - 0$, $t_2 - t_1$, ... are independent geometric random variables. Note that excursions

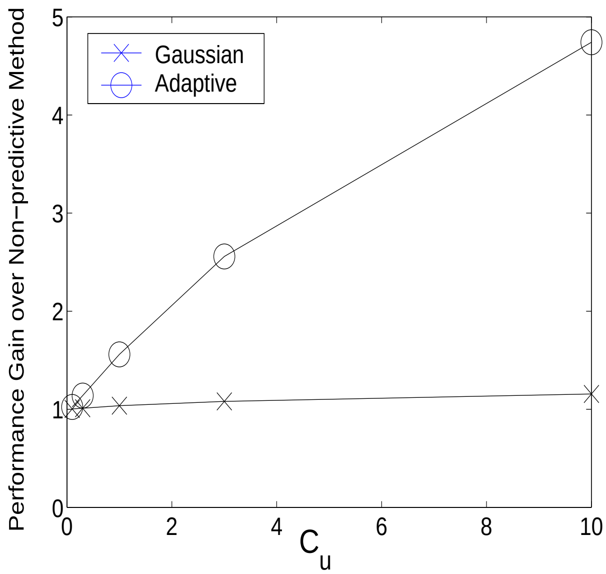

Figure 9: Performance gain as a function of the update cost C_u. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $m = 10$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.001 and 0.024. All differences between Gaussian and adaptive are statistically significant at the 1% level.

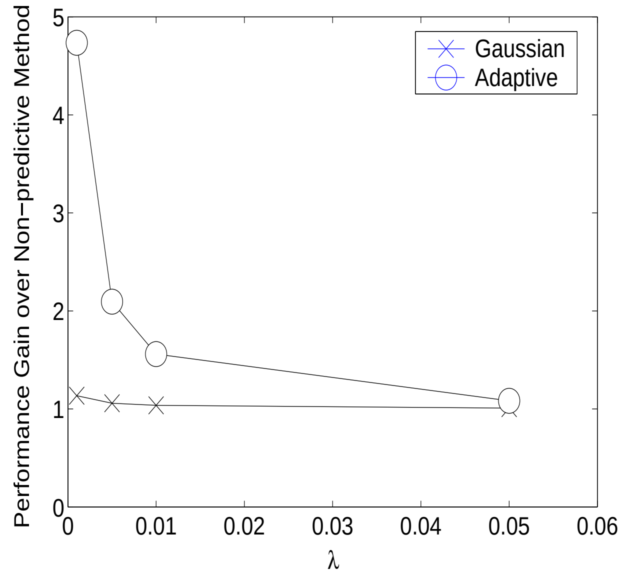

Figure 10: Performance gain as a function of the call frequency λ. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $m = 10$, and $C_u = 1.0$. Standard errors for the performance gains range between 0.001 and 0.051. All differences between Gaussian and adaptive are statistically significant at the 1% level.

begin at times $0, t_1, \ldots$. At each time t_i a *mean* velocity, denoted μ_i, for the upcoming excursion is chosen from a normal distribution with mean μ and variance denoted τ^2. Then, given μ_i, the *initial* velocity for that excursion, $v(t_i)$, is chosen from a normal distribution with mean μ_i and variance σ^2. The μ_i are chosen independently. Note that $v(t_i)$ is independent of $v(t_i - 1)$; that is, when the mobile reaches a destination, it chooses a new velocity independent of the previous one. Let $\varepsilon(t)$ denote an independent and identically distributed Gaussian process with mean 0 and variance σ^2. For any time t with $t_i < t < t_{i+1}$, the velocity at time t is given by

$$v(t) = \alpha v(t-1) + (1-\alpha)\mu_i + \sqrt{1-\alpha^2}\,\varepsilon(t) \tag{5}$$

Note that Equation (5) is identical to Equation (1), except that the mean velocity is μ_i rather than μ. In other words, the mean velocity changes every time a destination is reached, instead of being fixed indefinitely. Thus in this model, mobility is described by a sequence of independent Gauss-Markov processes whose lengths are geometrically distributed. We denote the renewal rate by λ_{renew}, so that the mean duration of an excursion is $1/\lambda_{renew}$.

We compared the performances of the Gaussian, adaptive, and non-predictive methods through simulation. As in section 3, we generated, for each simulation, a realization of our Gauss-Markov model with renewal over a time period that spanned 10,000 calls. There are eight parameters governing the performance. These include the six parameters discussed in section 3, plus λ_{renew} and τ.

Figures 11–14 present performance gains of the Gaussian and of the adaptive methods for a variety of choices of the parameters. The adaptive method outperforms the Gaussian method in many cases, and both outperform the non-predictive method. In general, the reason for the superior performance of the adaptive method is that it re-estimates the mean velocity at each update, thus adjusting to changing mobility conditions. In contrast, the Gaussian method always uses the asymptotic mean μ, while on any excursion the actual mean is μ_i. Thus, within any excursion, the Gaussian method misspecifies the mean velocity, but over many excursions, the misspecification errors will tend to cancel out.

Figure 11 plots performance gain vs. λ_{renew}, where λ_{renew} takes the values 0.001, 0.005, 0.01, 0.025, and 0.05. The figure shows that the superiority of the adaptive method is greater when the renewal rate is smaller, that is, when the times between destinations are longer. When excursion times are longer, the error in location estimation due to misspecification of the mean velocity accumulates to a greater extent, leading to more frequent updates, and to more paging when calls arrive. On the other hand, when excursions are short, it is more likely that several excursions, with differing μ_i, will occur between calls, or even between updates. In these cases, the misspecification errors caused by the differences between the μ_i and μ will tend to cancel out.

Figure 12 plots performance gain vs. τ, where τ takes the values 0.1, 0.25, 0.5, and 1.0. The parameter τ is the standard deviation of the excursion mean velocities μ_i, so it indicates the likely magnitude of the difference between any given μ_i and μ. Thus, when τ is small, μ_i tends to be close to μ, and the Gaussian method performs well. When the variation in the excursion mean velocities is equal to the variation in velocity within an excursion ($\tau = \sigma = 0.5$), the adaptive method is superior.

Figure 13 plots performance gain vs. α, where α takes the values 0, 0.5, 0.75, 0.9, and 0.99. The figure indicates that unless α is fairly large, the adaptive method outperforms the Gaussian method. Figure 14 plots performance gain vs. μ, where μ takes the values 0, 0.5, 1.0, and 5.0. The value of α is fixed at 0.75. The plot shows that the advantage of the adaptive method does not depend much on the asymptotic mean μ.

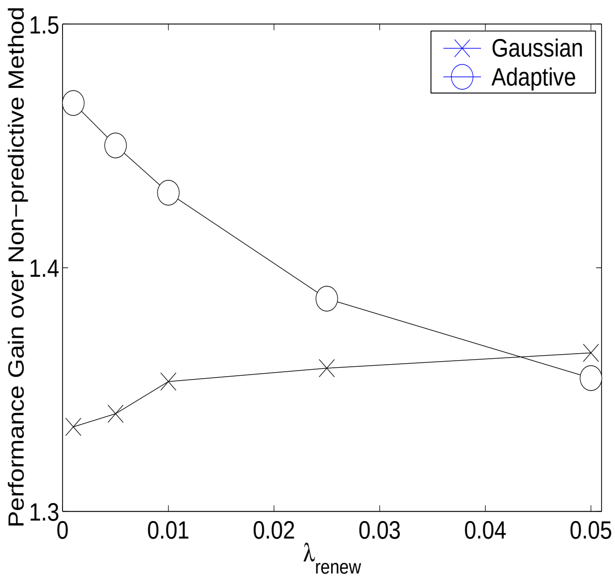

Figure 11: Performance gain as a function of the renewal rate λ_{renew}. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, $\lambda = 0.01$, and $\tau = 0.5$. Standard errors for the performance gains range between 0.006 and 0.022. The difference between Gaussian and adaptive at $\lambda_{renew} = 0.05$ is not statistically significant at the 5% level; all other differences are statistically significant at the 1% level.

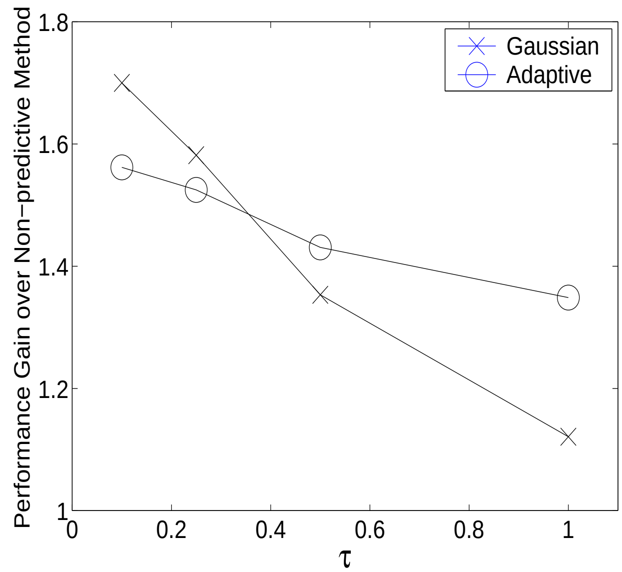

Figure 12: Performance gain as a function of τ. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, $\lambda = 0.01$, $\lambda_{renew} = 0.01$. Standard errors for the performance gains range between 0.004 and 0.011. All differences between Gaussian and adaptive are statistically significant at the 1% level.

6 Comparison when the Gauss-Markov model is misspecified

We now compare the Gaussian and the adaptive methods in situations where the Gauss-Markov mobility model (Equation 1) holds, but the parameters are misspecified. Thus the location updates in the Gaussian method are performed with inaccurate values of α or μ. Specifically, we will denote the true values of the correlation and asymptotic mean velocity that govern the mobility by α and μ respectively, but will assume that location updates are computed by using values α^* and μ^*, which may differ from α and μ. The adaptive method does not use parameters, so its location updates are unaffected.

We compared the performances of the Gaussian, adaptive, and non-predictive methods through

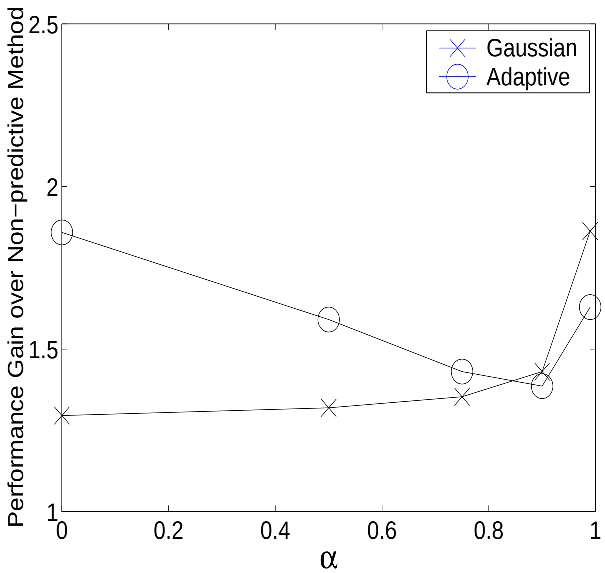

Figure 13: Performance gain as a function of the correlation α. Values of the other parameters are $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, $\lambda = 0.01$, $\lambda_{renew} = 0.01$, and $\tau = 0.5$. Standard errors for the performance gains range between 0.006 and 0.013. All differences between Gaussian and adaptive are statistically significant at the 1% level.

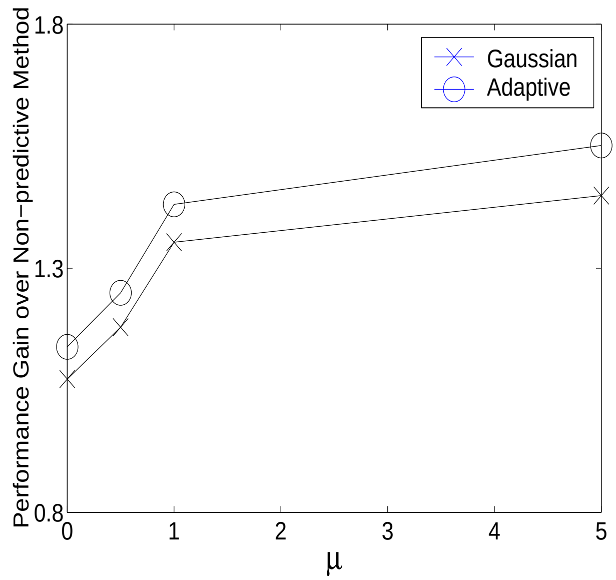

Figure 14: Performance gain as a function of the mean velocity μ. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $m = 10$, $C_u = 1.0$, $\lambda = 0.01$, $\lambda_{renew} = 0.01$, and $\tau = 0.5$. Standard errors for the performance gains range between 0.004 and 0.007. All differences between Gaussian and adaptive are statistically significant at the 1% level.

simulation. As in section 3, we generated, for each simulation, a realization of the Gauss-Markov model over a time period that spanned 10,000 calls. There are eight parameters governing the performance. These include the six parameters discussed in section 3, plus α^* and μ^*.

We first discuss the situation where μ is misspecified. We ran simulations with true values of μ taken from 0, 0.5, 1.0, and 5.0. For each of these values of μ we ran simulations with $\mu^* = \mu + 0.1$, $\mu^* = \mu + 0.5$, and $\mu^* = \mu + 1.0$. The correlation α was correctly specified in all of these simulations. Figure 15 presents performance gains of the adaptive method versus the Gaussian method. Gains greater than 1 favor the adaptive method. When μ is slightly misspecified (i.e. by 0.1), the Gaussian method retains its advantage over the adaptive method. When the misspecification is as little as 0.5, however, the adaptive method is better. Note that the performance gain does not vary much with the true value μ. This indicates that it is the absolute value of the misspecification, rather than the misspecification relative to the true value, that determines the performance.

We next compare the methods in situations where the correlation α is misspecified. Figure 16 presents performance gains of the adaptive method versus the Gaussian method, where α takes on the values 0, 0.5, 0.75, 0.9, and 0.99, and α^* takes on the values 0.75, 0.90, and 0.99. In all of the simulations reported in this figure, the asymptotic mean μ is correctly specified. Gains greater than 1 favor the adaptive method. Unlike misspecification of the asymptotic mean, correlation misspecification does not have a great impact on performance unless either the true correlation or the misspecified correlation is large. For example, if $\alpha^* = 0.75$, the Gaussian method retains its superiority over the adaptive method so long as $\alpha < 0.9$. As α approaches 1, the performance of the Gaussian method deteriorates rapidly. In fact, the plot shows that the Gaussian method is quite sensitive to correlation misspecification when the correlation is large. When $\alpha = 0.99$, but the value $\alpha^* = 0.90$ is used for location prediction, the performance of the Gaussian method is degraded to the point where it underperforms the adaptive method. When $\alpha = 0.90$, but the value $\alpha^* = 0.99$ is used for updating, the Gaussian method is essentially no better than the adaptive method.

7 Conclusions

We have introduced an adaptive method of mobility management for personal communications service networks in which the location of a mobile at the time of call arrival is predicted on the basis of information provided in past location updates, without the need to specify a mathematical model for mobility. Our method is simple, applicable to a wide variety of mobility patterns, and is particularly suited to conditions in which the mobility patterns are too complex to be described accurately by a tractable mathematical model. In contrast, methods that assume a parametric mobility model will perform well when their assumptions are met, but can perform much less well in other situations.

Simulation results show that the adaptive method outperforms a non-predictive method, in which paging always begins at the last updated location, under all mobility models we investigated except those where the true mean velocity is equal to 0. Furthermore, in a situation where the Gauss-Markov assumptions do not hold, the adaptive method outperforms the Gaussian method by a wide margin. In general, whenever the mean velocity of a mobile varies substantially over time, the adaptive method will outperform the Gaussian method. This is due to the fact that the adaptive method bases its estimate of mean velocity on a recently computed average velocity, rather than on a fixed value for a model parameter. In this way, the adaptive method adjusts to changing mobility conditions. In our simulations, the adaptive method outperformed the Gaussian method whenever

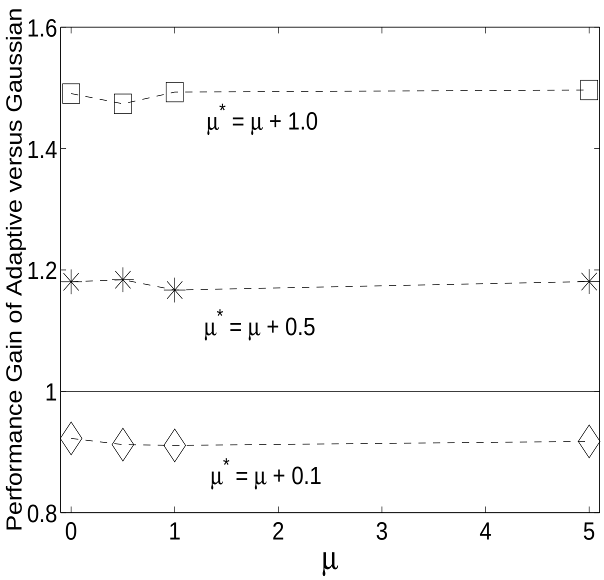

Figure 15: Performance gain of the adaptive method versus the Gaussian, as function of the mean velocity μ, when the value used for location prediction is μ^*. Gains > 1 favor the adaptive method. Values of the other parameters are $\alpha = 0.75$, $\sigma = 0.5$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.0033 and 0.0095. All performance gains are significantly different from 1 at the 1% level.

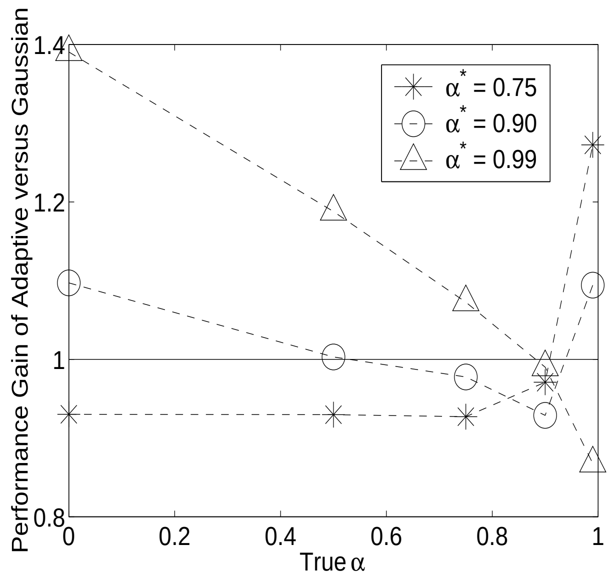

Figure 16: Performance gain of the adaptive method over the Gaussian, as a a function of the correlation α, when the value used for location prediction is α^*. Gains > 1 favor the adaptive method. Values of the other parameters are $\sigma = 0.5$, $\mu = 1.0$, $m = 10$, $C_u = 1.0$, and $\lambda = 0.01$. Standard errors for the performance gains range between 0.002 and 0.017. The performance gains for $(\alpha, \alpha^*) = (0.50, 0.90)$ and $(0.90, 0.99)$ are not significantly different from 1 at the 5% level. All others are significantly different from 1 at the 1% level.

the long-term variation in velocity is greater than or equal to the short-term variation (i.e. $\tau \geq \sigma$).

For the Gaussian method to perform well, the Gauss-Markov model must hold at least approximately. Accurate specification of the asymptotic mean μ is more important than accurate specification of the correlation α. This is intuitively plausible, since μ is the first moment of the velocity distribution, while α depends only on second moments. Even when the Gauss-Markov mobility model holds precisely, and the parameters are correctly specified, the performance of the adaptive method is close to that of the Gaussian method.

Parameters not related to mobility affect the performance of the adaptive method relative to other methods. As the cost of an update relative to the cost of a page (C_u) decreases, the performances of the Gaussian, adaptive, and non-predictive methods all converge. The same phenomenon is observed as the frequency of calls increases.

References

[1] Forman, G and Zahoran J (1994). The Challenges of Mobile Computing. *IEEE Computer*, 27(4):38–46.

[2] Liang, B and Haas, Z (1999). Predictive Distance-Based Mobility Management for PCS Networks. *Proceedings of the Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'99)*, New York, New York, USA, pp. 1377–1384.

[3] Wong, V and Leung, V (2000). Location management for next generation personal communications networks. *IEEE Network*, 14:18–24.

[4] Bar-Noy, A, Kessler, I, and Sidi, M (1995). Mobile users: To update or not to update? *ACM/Baltzer Wireless Networks: The Journal of Mobile Communication, Computation and Informatio* 1:175–195.

[5] Rose, C (1996). Minimizing the average cost of paging and registration: A timer-based method. *ACM/Baltzer Wireless Networks: The Journal of Mobile Communication, Computation and Information* 2:109–116.

[6] Brown, TX and Mohan S (1997). Mobility management for personal communication systems. *IEEE Transactions on Vehicular Technology* 46:269–278.

[7] Madhow, U, Honig, ML, and Steiglitz, K (1995). Optimization of wireless resources for personal communications mobility tracking. *IEEE/ACM Transactions on Networking*, 3:698–707.

[8] Wong, V and Leung, V (2001). An adaptive distance-based location update algorithm for next-generation PCS networks. *IEEE Journal on Selected Areas in Communications*, 19:1942–1952.

[9] Bhattacharya, A and Das, S (1999). LeZi-Update: An information-theoretic approach to tracking mobile users in PCS networks. *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'99)*, Seattle, WA, USA, pp. 1–12.

[10] Cayirici, E and Akyildiz, I (2002). User Mobility Pattern Scheme for Location Update and Paging in Wireless Systems *IEEE Transactions on Mobile Computing*, 1:236–247.

[11] Biesterfeld, J and Jobmann, K (1998). The use of prediction areas to improve mobility management algorithms. *Proceedings of the International Conference on Telecommunications '98*, Chaldiki, July 1998, **4**, pp. 446–450.

[12] Pathirana, P, Savkin A, and Jha, S (2003). Mobility modelling and trajectory prediction of cellular networks with mobile base stations. *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc'03)*, Annapolis, Maryland, USA, pp. 213–221.

[13] Zomaya, A, Haydock M, and Olariu, S (2003). Some observations on using meta-heuristics for efficient location management in mobile computing networks. *Journal of Parallel and Distributed Computing* **63**, 33–44.

[14] Hwang, S, Han, Y, Lee, B, and Hwang C (2003). An adaptive location management scheme using the velocity of nodes. *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'03)*, New Orleans, Louisiana, USA, pp. 1999–2004.