

CEAS TECHNICAL REPORT NO. 660

ver. January 26, 1993

NOTE: This is an updated and revised version of
CEAS Technical Report No. 610.

**TRAFFIC PERFORMANCE FOR CELLULAR COMMUNICATION SYSTEMS
WITH MIXED PLATFORMS AND IMPERFECT HAND-OFF INITIATION**

by

Stephen S. Rappaport
Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794-2350
USA

phone: (516) 632-8394
fax: (516) 632-8494
e-mail: rappap@sbee.sunysb.edu

KEYWORDS:

cellular communication networks, mobile communications, personal communication networks,
radiotelephone, telecommunications traffic, wireless networks, blocking, hand-off,
hand-over.

ABSTRACT

Cellular communication systems that support a mixture of platform types distinguished by different mobility characteristics are considered. A tractable analytical model for traffic performance analysis is developed using multi-dimensional birth-death processes and the method of phases. The framework allows consideration of homogeneous and non-homogeneous systems, a broad class of dwell time distributions, and "missed" hand-off initiations. Cut-off priority for hand-offs and several platform types are considered to demonstrate the approach. The effects of different mobility parameters and of imperfect detection of hand-off needs are examined. Theoretical performance characteristics are obtained. These exhibit carried traffic, hand-off activity, blocking probability and forced termination probability for each platform type. The realizable exchange of blocking for hand-off performance is shown. Computational issues are discussed and a method for extrapolating performance characteristics for systems with many channels is presented.

The author is grateful to Mr. Gustavo MONTE for skillful computer programming that generated the performance characteristics shown.

The research reported in this paper was supported in part by the U.S. National Science Foundation under Grant No. NCR-9025131 and in part by IST/SDIO under Grant No. N00014-91-J-4063 administered by the U.S. Office of Naval Research. Some of the results were obtained using the Cornell National Supercomputer Facility, a resource of the Center for Theory and Simulations in Science and Engineering (Cornell Theory Center), which receives major funding from the U.S. National Science Foundation and IBM Corporation, with additional support from New York State and members of the Corporate Research Institute.

INTRODUCTION

The hand-off problem arises in cellular mobile and personal communication networks when a communicating platform moves from a spatial (source) region served by one wireless gateway to a (target) region served by another. When circuit or virtual circuit switching is used, continuation of a call depends on several factors. These include: (1) reliable detection of conditions that indicate the need of a platform for a hand-off; (2) identification of alternative gateway(s) for service; (3) timely exchange of supervisory signals; and, (4) deployment and availability of communications resources. Missed or unsuccessful hand-offs result in forced termination of calls and are perceived by users as interruptions in service. As smaller cell sizes are used to increase overall system capacity, more hand-offs are needed to sustain calls to satisfactory completion. Thus the problem is becoming increasingly significant as mobile, portable and personal communications proliferate and small-cell high-capacity systems are devised.

Early studies of hand-off issues relied exclusively on computer simulation [1]. Development of realistic analytically tractable models has been the thrust of some more recent work [2]-[8]. Among these is a methodology, which is based on the notion of multi-dimensional birth-death processes [6]-[8]. It permits the computation of theoretical performance characteristics, provides considerable additional insight into the problem, and alleviates the sole dependence on simulation techniques - even for situations involving a fair amount of physical complexity. The method depends heavily on the assumption of memoryless properties for various driving processes [9]. Similar assumptions in other contexts have served telecommunications traffic theory well for many years, but since less is known about some of the driving processes that arise in the cellular communications context, it is important to extend the applicability of the approach.

This paper extends our earlier results and mathematical framework in several important directions. Firstly, we present a state variable description that allows consideration of a broad class of dwell time distributions which can accommodate empirical data, while at the same time retaining the conditions that allow an analytically tractable solution. Secondly, a mixture of platform types (with widely differing mobility characteristics) in the same system is considered. Thirdly, the analytical model is extended to include the effect of imperfect hand-off initiation. Specifically we develop formulas that relate overall hand-off performance to two component problems - *hand-off initiation* and *resource availability*. Finally, we describe our computational procedure and present example results obtained via workstation and remote access to supercomputing facilities.

In previous work we used the concept of *dwell time* to characterize the amount of time that a mobile platform is *within communications range* of a given gateway. Figure CS:1.1 shows an *approximate* relationship between cell sizes and dwell times for various platform types. Estimated speeds of the platform types listed are very rough, and of course, there is no definitive agreement on terminology for cell sizes. The figure is NOT intended to suggest, for example, that jetliners and pedestrians should be supported by the same system and on the same channels. Indeed, strong arguments can be made which suggest that even hand-held and automobile mounted units should be serviced by *different* wireless systems with appropriate interconnecting gateways [10],[11],[12]. But the figure DOES illustrate, that for a given cell size and system, there is a variety of platform types (with a range of mobili-

ty characteristics) that should be supported. Moreover, there exists for example, the possibility that hand-held units (intended for pedestrians) will frequently be used on more highly mobile platforms such as an automobiles. Such use can have a significant impact on communications traffic performance. Thus the consideration of mixed platform types is an issue that should be considered. The analytical basis for Figure CS:1.1 is given in [13].

Figure CS:1.1 serves as a useful *starting point* but it is in fact only that. It is important to understand the meaning of dwell time in the context of the physical situation. The definition of dwell time which we present above is broad enough to include important effects of the radio environment, such as fading. One need only realize that the term *within communications range* used in the definition implies that an acceptable quality (two-way) communication link can be established. Thus *range* is not necessarily a fixed distance. The dwell time is determined by the duration of time that such a two-way link can be maintained, for whatever reasons. These reasons include fading and propagation effects, in addition to path loss (which is largely determined by distance between mobile and gateway).

Because fading phenomena can affect signal quality even when a platform is not moving, one could consider that even a mobile platform that is stationary will have some (random) finite dwell time. Other scatterers and objects in the environs of the gateway and platform can and do move quite indepen- dently of any given mobile platform - and these scatterers affect signal quality for the mobile of interest. Thus, while it may seem at first glance that the framework under discussion here does not consider the radio environ- ment, SOME ASPECTS ARE INCLUDED. It is the detailed modeling of the radio environment that is omitted or circumvented. The overall effect of the radio environment and the mobility of a platform are included in the notion of a random dwell time. The omission of detail allows a relatively simple repre- sentation. In principle, dwell time distributions can be easily measured once suitable definitions of acceptable signal quality are specified. The approach that is developed here would allow such empirically determined parameters of the dwell time distribution to be incorporated into overall performance char- acteristics. Such characteristics would also reflect communication traffic and resource oriented aspects of the problem which are the main focus of this development. Moreover, in subsequent sections we consider the sensitivity of the overall performance characteristics to the dwell time distribution. In particular we find that the major performance characteristics such as blocking probability, forced termination probability, and carried traffic depend pri- marily on mean dwell time and are insensitive to its variance. In this way, the analysis provides important clues (with justification) about *what observa- tions should be made* in an empirical validation of a proposed channel assign- ment strategy, and *what parameters should be extracted* from the observations.

Nevertheless, the model presented here does not specifically consider the *dropping* of calls due to fading. This aspect of the problem is deferred for subsequent investigation. Inclusion of such effects would require some further extension (and additional parameters). The main thrust of our work has been to develop a framework for consideration of alternative hand-off strategies. Consequently the performance characteristics that result may be optimistic in this sense for some applications. But fading is much more of a problem for large-cell systems than for small-cell systems - especially those for which a line-of-sight path between mobile and base (gateway) is likely. And, it is precisely for small-cell systems that the hand-off problem is particularly acute because of the larger numbers of cell boundary crossings. Thus the omission of call dropping due to fading, and emphasis on the impact

of hand-off strategies is quite reasonable (and probably not optimistic) for such systems.

Clearly, dwell time depends on many factors such as propagation conditions, the path a mobile platform follows, its velocity profile along the path, and especially the definition of *communications range*. But even if all of these were known, the dependency is so complex and burdensome that one would eventually have to resort to empirical findings in some way. Much of the difficulty can be circumvented by characterizing dwell time as a random variable whose distribution function is known but chosen to accommodate empirical data.

In the simplest of the models that fit into the framework, for each platform type, one can take the probability density function (pdf) of dwell time as a negative exponential function having a parameter chosen to accommodate estimates or measurements. Since this pdf has the memoryless property one of the necessary conditions for characterization of state transitions as a multi-dimensional birth-death process is satisfied. Because of the myriads of paths and velocity profiles that individual platforms can follow, in the absence of contrary empirical data, this form of pdf is probably as good as any other. It also provides a significant advantage from the mathematical viewpoint. One characteristic of the negative exponential distribution (ned), is that it is weighted toward low values. That is, for a ned variate, about 63% (fraction = $1 - e^{-1}$) of realizations are less than the mean. For a pdf that is symmetrical about the mean, such as the Gaussian pdf, of course, only 50% of the realizations are below the mean. So it is interesting to consider how appropriate this model is in the light of physical situations that occur in the mobile radio context.

The coefficient of variation for a random variable is defined as the ratio of its standard deviation to its mean. Roughly, it is a measure of the dispersion of the random variable about its central value. Higher values indicate greater relative dispersion. For an exponential variate this coefficient is unity (a relatively high value). So if platforms follow many differing paths and widely varying speed profiles, a highly dispersed variate such as one with a ned is reasonable in this regard. In actuality, many cells have *holes*. These are areas that are completely surrounded by the cell but which are out of range of the gateway serving that cell. Thus, if a communicating platform in a cell moves into a hole, it must be handed off to the gateway serving the hole or the call will be terminated. In any case, platforms that enter holes are considered to leave the cell. Of course many platforms do not enter holes and therefore do not leave the cell until reaching the "more conventional" boundary. Because of this phenomenon too, dwell times would (with some randomness) tend to be shorter than the cell geometry would otherwise suggest. So the distribution would be skewed toward the lower values as in the case for the negative exponential.

On the other hand, if a significant fraction of platforms that enter and leave a cell are part of vehicular highway traffic streams, the situation is different. Such platforms tend to move at a more constant speed, along essentially fixed paths, and within regions of good radio coverage where there are few holes, if any. These platforms would tend to have dwell times that are more highly concentrated about some mean. The pdf would be characterized by a coefficient of variation that is less than unity. Therefore we are motivated to generalize the family of dwell time pdf's that can be accommodated within the framework that we are developing. One way to do this was de-

scribed recently although a full discussion of the generalization was not given [7],[8]. The detailed development of this important extension is presented here along with sample theoretical performance characteristics.

The approach used is to consider the dwell time to be a random variable that is the sum of independent negative exponential random variables. Then by using an appropriate definition of state variables, the memoryless property can be maintained, and a multi-dimensional birth-death process model can be used. Dwell times are considered to consist of several phases. This artifice permits consideration of pdf's that are not memoryless within the same framework. A similar generalization for unencumbered session times is also possible but is not developed here for two reasons. Firstly, the approach tends to increase the number of states needed to represent the system. Secondly, session duration times that have negative exponential distributions are not uncommon in telecommunications traffic theory, so there are ample precedents. We focus on dwell time characteristics, because much less is known about these, and as explained above it seems that a broad class of distributions that characterize dwell times would be useful in modeling actual cellular systems.

Theoretical performance characteristics are calculated and presented. These show carried traffic, hand-off activity, blocking probability and forced termination probability for each platform type. The characteristics exhibit the realizable exchange of blocking for hand-off performance. The effects of different platform mobility parameters on performance are investigated. *Extended performance measures* which account for: (1) the ability of the detection/decision algorithm to initiate the hand-off procedure in sufficient time to allow the required exchange of supervisory signals, as well as, (2) the availability of communications resources in the target cell, are developed. The approach and results are applicable to cellular and micro-cellular systems. For some parameter choices, calculations to generate the results were done on readily available workstations. For others, where a large number of states were required, calculations were done using remote access to the Cornell National Supercomputer Facility.

MODEL DESCRIPTION

Since the development presented here proceeds along the lines of [6]-[8], we present only a brief description of the model and then proceed with the formulation that allows generalization of the dwell time distribution. We consider a large geographical region covered by cells, that are defined by proximity to designated network gateways. The region is traversed by large numbers of mobile platforms that are of several types. In the development presented here, a platform can support at most one call. The platform types differ primarily in their mobility characteristics. Pedestrians with hand-held devices and autos with cellular phones are example platform types. Communication with a mobile platform is via a wireless gateway node or nodes identified with each cell. We use the word *cell* in its generic sense to describe a spatial region serviced by a wireless gateway. The essential problem considered here is the same whether this is a macro-cell, micro-cell, pico-cell, zone, sector, or satellite beam. Thus the particular configuration is not very important in terms of demonstrating the applicability of the approach. However, Figure CS:1.2, which some readers may find helpful, suggests an example configuration to which the analysis applies.

CHANNEL LIMIT: Each cell or gateway can accommodate C channels.

CHANNEL QUOTAS: At any gateway, the maximum number of channels that can be simultaneously used by g-type platforms is J(g).

CUT-OFF PRIORITY: C_h channels in each cell are reserved for hand-off calls. New calls will be blocked if the number of channels in use is $C - C_h$ or greater. Hand-off attempts will fail if the number of channels in use is C.

The unencumbered call (session) duration on a g-type platform is a ned random variable, $T(g)$, having a mean $\bar{T}(g) = 1/\mu(g)$.

The dwell time in a cell for a g-type platform is a random variable, $T_D(g)$ having a mean $\bar{T}_D(g)$. The random variable, $T_D(g)$ is the sum of $N(g)$ statistically independent ned random variables denoted $T_D(g,i)$, where $i=1,2,3,\dots,N(g)$. The mean of $T_D(g,i)$ is $\bar{T}_D(g,i) = 1/\mu_D(g,i)$ and its variance is $VAR(T_D(g,i)) = 1/[\mu_D(g,i)]^2$. Thus

$$\bar{T}_D(g) = \sum_{i=1}^{N(g)} \bar{T}_D(g,i) = \sum_{i=1}^{N(g)} 1/\mu_D(g,i) \quad , \quad (1)$$

and

$$VAR[T_D(g)] = \sum_{i=1}^{N(g)} 1/[\mu_D(g,i)]^2 \quad . \quad (2)$$

The squared coefficient of variation for the dwell time of a g-type platform is the ratio of its variance to the square of its mean. This is given by

$$\kappa^2[T_D(g)] = VAR[T_D(g)] / [\bar{T}_D(g)]^2 \quad . \quad (3)$$

An interesting special case occurs if $\mu_D(g,i)$ does not depend on i . Then $T_D(g)$ has an Erlang distribution with a mean $N(g)/\mu_D(g)$, and a squared coefficient of variation, $1/N(g)$. For each platform type, the two parameters can be easily chosen to accommodate empirical data by using the mean and variance of observed dwell times. A value of $\kappa^2=1$ corresponds to a ned random variable, while $\kappa^2=0$ corresponds to the deterministic case (constant dwell time).

The problem is to calculate relevant performance characteristics, including, (for each platform type): Blocking Probability, Hand-off Failure Probability, Forced Termination Probability, Carried Traffic, and Hand-Off Activity.

Notes: We consider blocking probability to be the average fraction of *new* call originations that are denied access to a channel. Hand-off failure probability is the average fraction of hand-off "needs" that fail to gain access to a channel in the target zone. Forced termination probability is the probability that a call will suffer a hand-off *attempt* failure some time in the "lifetime" of the call. Hand-Off Activity is the average number of hand-off attempts for a call that receives service.

The mathematical analysis is similar to that used in [6] - [8]. The major differences are in definition of the state variables, identification of

the driving processes, and the equations that specify the state probability transition flows. In what follows we emphasize those aspects of the mathematical development that differ. We consider that when a platform of type g enters a cell, it passes through $N(g)$ phases of dwell time. These are identified by the index, $i=1,2,3,\dots,N(g)$. The amount of time spent in each phase is $T_D(g,i)$, a ned random variable as defined above.

STATE CHARACTERIZATION

First consider a *single* cell. We define the state (of a cell) by a sequence of non-negative integers. These can be conveniently written as G n-tuples.

$$\begin{array}{cccccccc}
 v_{11}' & v_{12}' & v_{13}' & \cdots & v_{1N(1)} & & & \\
 v_{21}' & v_{22}' & v_{23}' & \cdots & \cdots & \cdots & v_{2N(2)} & \\
 : & : & : & & & & : & \\
 v_{g1}' & v_{g2}' & v_{g3}' & \cdots & \cdots & \cdots & v_{gN(g)} & \\
 : & : & : & & & & : & \\
 v_{G1}' & v_{G2}' & v_{G3}' & \cdots & \cdots & \cdots & v_{GN(G)} &
 \end{array} \quad (4)$$

where v_{gi} { $g=1,2,\dots,G$; $i=1,2,\dots,N(g)$ } is the number of platforms of type g that are in phase i . It was found convenient to order the states using an index $s=0,1,2,\dots,s_{\max}$. Then the state variables v_{gi} can be shown explicitly dependent on the state. That is, $v_{gi} = v(s,g,i)$.

When the cell (gateway) is in state, s , the following characteristics can be determined:

The number of channels being used by g -type platforms is

$$j(s,g) = \sum_{i=1}^{N(g)} v(s,g,i) \quad . \quad (5)$$

The total number of channels in use is

$$j(s) = \sum_{g=1}^G j(s,g) \quad . \quad (6)$$

Permissible states correspond to those sequences for which all constraints are met. Additional constraints can also be considered within this same framework. Here we have a *channel limit* which requires, $j(s) \leq C$, and *channel quotas* which require, $j(s,g) \leq J(g)$, for $g=1,2,\dots,N(g)$.

A thorough formulation which accounts for direct coupling of *cell state* transitions of adjacent cells that are involved in a hand-off is circumvented by relating the average hand-off arrival and departure rates as in [6] and [7]. This avoids having to deal with an enormously (and usually intractably) larger number of *system states* represented by sequences of all simultaneously possible *cell states*. Both *homogeneous* and *non-homogeneous* cellular systems can be treated in this way. The result is that we only have to consider a *single cell* and deal with the number of states needed to characterize its behavior. Even with this simplification, the number of states that are needed can be quite formidable for certain parameter choices. So, it is important

that the procedures and algorithms used to solve the resulting equations be chosen, devised, and organized so that computational solutions are feasible. The number of cell states needed to represent a system with two platform types each having two phases of dwell time and no channel quotas is shown in Table 1.

TABLE 1: NUMBER OF CELL STATES NEEDED
 $G=2, N(1)=2, N(2)=2, \text{CHANNEL LIMIT} = C.$

C	number of states
10	1,001
15	3,876
20	10,626
25	23,751
30	46,376
36	91,390

DRIVING PROCESSES AND STATE TRANSITION FLOWS

For this example there are five relevant driving processes. These are: {n} The generation of new calls in the cell of interest; {c} The completion of calls in the cell of interest; {h} The arrival of communicating vehicles at the cell of interest; {d} The departure of communicating vehicles from the cell of interest; and, {φ} The transition between dwell time phases. Because there are different platform types, all of these processes are multi-dimensional. As before we use Markovian assumptions for the driving processes in order to render the problem amenable to solution using multi-dimensional birth-death processes. Specifically, in addition to the previous assumptions we assume that: {1} The new call arrival processes in any state are Poisson point processes with state dependent means; and that, {2} The hand-off call arrivals for each platform type are Poisson point processes.

It remains to characterize all of the state transitions. For each state, the possible *predecessor* states must be identified. That is, those states which could have *immediately* given rise to the current state under each of the (multi-dimensional) driving processes. In addition, the state probability transition flows, must be found. The flow balance equations can then be determined and the state probabilities can be calculated [6],[9].

In subsequent paragraphs, two dummy indices, z_1 and z_2 , are used for convenience as follows: $z_1=1,2,3,\dots,G$; $z_2=1,2,3,\dots,N(z_1)$.

New Call Arrivals

A transition into state s due to a new call arrival on a g -type platform in phase i when the cell is in state x_n , will cause the state variable $v(x_n, g, i)$ to be incremented by 1. We note that because cut-off priority is used, C_h channels are held for hand-off arrivals. A new call can be served only if the number of channels in use does not exceed $C - C_h$. Thus a permissible state x_n is a predecessor state of s for new call arrivals on g -type platforms in phase i , if $j(x_n) < C - C_h$, $j(x_n) < J(g)$, and the state variables are related by

$$v(x_n, g, i) = v(s, g, i) - 1$$

$$v(x_n, z_1, z_2) = v(s, z_1, z_2) \quad , \quad z_1 \neq g \quad , \quad (7)$$

$$v(x_n, z_1, z_2) = v(s, z_1, z_2) \quad , \quad z_2 \neq i \quad .$$

Let $\Lambda_n(g, i)$ denote the average arrival rate per cell of new calls from g-type platforms in phase i of dwell time. Under the assumption that a new call on a g-type platform is equally likely to arrive at any time during the platform's sojourn in a cell, the fraction of new call arrivals in a cell that arise from g-type platforms in phase i is

$$\rho_n(g, i) = \bar{T}_D(g, i) / \bar{T}_D(g) \quad . \quad (8)$$

Thus,

$$\Lambda_n(g, i) = \rho_n(g, i) \cdot \Lambda_n(g) \quad . \quad (9)$$

The flow into state s from x_n due to new call arrivals is $\gamma_n(s, x_n) = \Lambda_n(g, i)$. This can be expressed as

$$\gamma_n(s, x_n) = \rho_n(g, i) \cdot \alpha(g) \cdot \Lambda_n(1) \quad . \quad (10)$$

If all phases of the same type are statistically identical, then $T_D(g)$ has an Erlang distribution and $\rho_n(g, i) = 1/N(g)$.

Call Completions

A transition into state s due to a call completion on a g-type platform in phase i when the cell is in state x_c , will cause the state variable $v(x_c, g, i)$ to be decreased by 1. Thus a permissible state x_c is a predecessor state of s for call completions on g-type platforms in phase i, if the state variables are related by

$$\begin{aligned} v(x_c, g, i) &= v(s, g, i) + 1 \\ v(x_c, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_1 \neq g \quad , \quad (11) \\ v(x_c, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_2 \neq i \quad . \end{aligned}$$

The corresponding transition flow is given by

$$\gamma_c(s, x_c) = \mu(g) \cdot v(x_c, g, i) \quad . \quad (12)$$

Hand-Off Arrivals

Let Λ_h be the average rate at which hand-off arrivals impinge on the cell, and F_g denote the fraction of arrivals that are g-type platforms. Initially we assume that these parameters are known but ultimately we compute their values as part of the solution algorithm. Since a hand-off arrival is always due to a platform entering a cell, it corresponds to a platform in phase 1. Therefore, a transition into state s due to a hand-off arrival of a g-type platform when the cell is in state x_h , will cause the state variable $v(x_h, g, 1)$ to be incremented by 1. We also recall that hand-off arrivals have access to all C channels in a cell. Thus a permissible state x_h is a predecessor state of s for hand-off arrivals on g-type platforms, if $j(x_h) < C$, $j(x_h) < J(g)$, and the state variables are related by

$$v(x_h, g, 1) = v(s, g, 1) - 1 \quad ,$$

$$v(x_h, z_1, z_2) = v(s, z_1, z_2) \quad , \quad z_1 \neq g \quad , \quad (13)$$

$$v(x_h, z_1, z_2) = v(s, z_1, z_2) \quad , \quad z_2 \neq 1 \quad .$$

The corresponding transition flow is given by

$$\gamma_h(s, x_h) = \Lambda_h F_g \quad . \quad (14)$$

Hand-Off Departures

A hand-off departure corresponds to a platform completing its last phase. Thus a transition into state s due to a hand-off departure of a g -type platform when the cell is in state x_d , will cause the state variable $v(x_d, g, N(g))$ to be decreased by 1. Thus a permissible state x_d is a predecessor state of s for hand-off departures of g -type platforms if the state variables are related by

$$\begin{aligned} v(x_d, g, N(g)) &= v(s, g, N(g)) + 1 \\ v(x_d, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_1 \neq g \quad , \\ v(x_d, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_2 \neq N(g) \quad . \end{aligned} \quad (15)$$

The corresponding transition flow is given by

$$\gamma_d(s, x_d) = \mu_D(g) \cdot v(x_d, g, N(g)) \quad . \quad (16)$$

Dwell Time Phase Transitions

A transition into state s due to completion of a dwell time phase of a g -type platform in phase i when the cell is in state x_ϕ , will cause changes in two state variables simultaneously. The state variable $v(x_\phi, g, i)$ will be decreased by 1 and the state variable $v(x_\phi, g, i+1)$ will be increased by 1. This corresponds to a phase transition from i to $i+1$. We limit discussion in this paragraph to $1 \leq i < N(g)$ because completion of a dwell time phase $i=N(g)$ corresponds to a hand-off departure and was considered in a previous paragraph. Thus a permissible state x_ϕ is a predecessor state of s for phase transitions of g -type platforms in phase i , if the state variables are related by

$$\begin{aligned} v(x_\phi, g, i+1) &= v(s, g, i+1) - 1 \\ v(x_\phi, g, i) &= v(s, g, i) + 1 \\ v(x_\phi, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_1 \neq g \\ v(x_\phi, z_1, z_2) &= v(s, z_1, z_2) \quad , \quad z_2 \neq i \end{aligned} \quad (17)$$

where $g=1,2,3,\dots,G$; and $i=1,2,3,\dots,N(g)-1$. The corresponding transition flow is given by

$$\gamma_\phi(s, x_\phi) = \mu_D(g) \cdot v(x_\phi, g, i) \quad . \quad (18)$$

FLOW BALANCE EQUATIONS

From the equations given above the total transition flow into s from any permissible state x can be found using

$$q(s,x) = \gamma_n(s,x) + \gamma_c(s,x) + \gamma_h(s,x) + \gamma_d(s,x) + \gamma_\phi(s,x) \quad , \quad (19)$$

in which $s \neq x$ and flow into a state has been taken as a positive quantity.

The total flow out of state s , is denoted $q(s,s)$ and is given by

$$q(s,s) = - \sum_{\substack{k=0 \\ k \neq s}}^{s_{\max}} q(k,s) \quad . \quad (20)$$

To find the statistical equilibrium state probabilities for a cell, we write the flow balance equations for the states. These are a set of $s_{\max}+1$ simultaneous equations for the unknown state probabilities, $p(s)$. They are of the form

$$\sum_{j=0}^{s_{\max}} q(i,j) p(j) = 0 \quad , \quad i = 0,1,2,\dots,s_{\max}-1 \quad (21)$$

$$\sum_{j=0}^{s_{\max}} p(j) = 1 \quad ,$$

in which, for $i \neq j$, $q(i,j)$ represents the net transition flow into state i from state j , and $q(i,i)$ is the total transition flow out of state i . These equations express that in statistical equilibrium, the net probability flow into any state is zero and the sum of the probabilities is unity. The index, i , in (21) can run up to s_{\max} to provide a redundant set that may be helpful in numerical computation.

DETERMINATION OF THE HAND-OFF ARRIVAL PARAMETERS

In the foregoing analysis, it was assumed that the average hand-off arrival event rate, Λ_h , and the fractions of hand-off arrivals that are g -type platforms, that is, F_g , ($g=1,2,\dots,G$) are given. Actually these parameters are implicit, and must be determined from the dynamics of the process itself. The iterative method described in [6] can be used. Specifically, let $d_{g|s}$, denote the probability that a hand-off departure of a g -type platform occurs when the cell is in state, s . Then,

$$d_{g|s} = \mu_D(g) \cdot v(s,g,N(g)) / r(s) \quad . \quad (22)$$

Let R be the total average rate at which transition events occur in the cell.

This is given by

$$R = \sum_{s=0}^{s_{\max}} r(s) \cdot p(s) \quad . \quad (23)$$

The overall probability of a g-type hand-off departure can be expressed as

$$d_g = \sum_{s=0}^{s_{\max}} d_{g|s} \cdot \bar{p}(s) \quad , \quad (24)$$

in which, for $s=0, \dots, s_{\max} - 1$, $\bar{p}(s)$ is the probability that the system visits state s . The probability, $\bar{p}(s)$, is the state probability of the "jump" Markov chain [18] associated with the multi-dimensional birth-death process described above and is given by $\bar{p}(s) = [p(s) \cdot r(s)] / R$.

The fraction of hand-off departures that are g-type platforms is

$$F'_g = d_g / \sum_{g=1}^G d_g \quad . \quad (25)$$

The average hand-off departure rate can be expressed as

$$\Delta_h = R \cdot \sum_{g=1}^G d_g \quad (26)$$

We note that any hand-off departure of a g-type platform from a cell corresponds to a hand-off arrival of (the very same) g-type platform to some other cell. Thus, it must be true that for a homogeneous system in statistical equilibrium, the hand-off arrival and departure rates per cell must be equal. That is we must have

$$F_g = F'_g \quad , \quad (27)$$

and

$$\Lambda_h = \Delta_h \quad . \quad (28)$$

COMPUTATIONAL PROCEDURE

An important consideration is that the quantities, F'_g , Λ_h , and Δ_h which are reflected in the determination of the $q(i,j)$'s in the flow balance equations depend on the unknown state probabilities, $p(s)$. Therefore the flow balance equations are actually a set of simultaneous nonlinear equations. However, by beginning with guesses for Λ_h and F'_g , ($g=1,2,3,\dots,G$), the iterative approach of [6] can again be used. In this approach the system is linear within each iteration. Details of the computational procedure are described in the Appendix. The basic idea is to define a vector function, Q_1 , that (for a given Λ_h) determines the resulting F'_g 's and Δ_h , and another scalar

function

$$Q_2(\Lambda_h) = \Lambda_h - \Delta_h \quad . \quad (29)$$

The root of $Q_2(\Lambda_h) = 0$, gives the value of Λ_h that satisfies (28). Non-homogeneous systems can be treated similarly except that hand-off arrival and departure rates are related by

$$\Lambda_h = \theta \cdot \Delta_h \quad , \quad (30)$$

in which $\theta \neq 1$, is a "tilting parameter" which characterizes the non-homogeneity of the cell. For the non-homogeneous case the function Q_2 should be defined as

$$Q_2(\Lambda_h) = \Lambda_h - \theta \cdot \Delta_h \quad . \quad (31)$$

The overall procedure used three levels of nested iterations.

PERFORMANCE MEASURES

When the statistical equilibrium state probabilities and transition flows are found the required performance measures can be calculated.

Carried Traffic

The carried traffic per cell for each platform type is the average number of channels occupied by the calls from the given platform type. If users pay for "air time" this indicates the revenue that the service provider can expect from each type of platform. The carried traffic for g-type platforms is

$$A_C(g) = \sum_{s=0}^{s_{\max}} j(s,g) \cdot p(s) \quad , \quad (32)$$

and the total carried traffic is

$$A_C = \sum_{g=1}^G A_C(g) \quad . \quad (33)$$

Blocking Probability

The blocking probability for a call from a g-type platform is the average fraction of new g-type calls that are denied access to a channel. Blocking of new g-type call occurs if there are no channels to serve the call or if the number of g-type calls in progress is at the quota level. We define the following disjoint sets of states:

$$\begin{aligned} B_0 &= \{ s : C - C_h \leq j(s) \leq C \} \\ B_g &= \{ s : j(s) < C - C_h , j(s,g) = J(g) \} , \end{aligned} \quad (34)$$

in which, $g = 1, 2, \dots, G$.

Then the blocking probability for g-type calls is

$$P_B(g) = \sum_{s \in B_0} p(s) + \sum_{s \in B_g} p(s) \quad . \quad (35)$$

If there are no channel quotas, then $J(g) \geq C$, so there are no states in any of the sets, B_g , $g=1,2,3,\dots,G$. Blocking probability is then the same for any new call regardless of the platform type on which it originates.

Hand-Off Failure Probability

The hand-off failure probability for g-type calls is the average fraction of g-type hand-off attempts that are denied a channel. We note that hand-off attempts have potential access to all channels of a cell without regard to C_h , but may be subject to channel quota constraints. We define the following disjoint sets of states, in which at least some hand-off attempts will fail:

$$\begin{aligned} H_0 &= \{ s: j(s) = C \} \\ H_g &= \{ s: j(s) < C, j(s,g) = J(g) \} \quad , \end{aligned} \quad (36)$$

in which $g=1,2,3,\dots,G$.

Then the hand-off failure probability for a g-type hand-off can be written as

$$P_H(g) = \sum_{s \in H_0} p(s) + \sum_{s \in H_g} p(s) \quad . \quad (37)$$

Forced Termination Probability

Perhaps more important than hand-off failure probability, is the *forced termination probability*, $P_{FT}(g)$. This is defined as the probability that a g-type call that is not blocked is interrupted due to hand-off failure during its lifetime. For convenience we limit our discussion here to the case where all dwell-time phases of a given platform type are statistically identical. That is, $\mu_D(g,i) \equiv \mu_D(g)$. Then $\rho_n(g,i) = 1/N(g)$. Let $\pi(g)$ be the probability that a platform of type g completes its current dwell-time phase before its call is (satisfactorily) completed. Then since dwell time phases are independent random variables, we have

$$\pi(g) = \mu_D(g) / [\mu(g) + \mu_D(g)] \quad . \quad (38)$$

For a call that is being served on a g-type platform in phase i, $N(g)-i+1$ dwell time phases must be completed by the supporting platform for a hand-off attempt to be generated. Therefore the probability that such a call requires a hand-off is

$$b(g,i) = [\pi(g)]^{N(g)-i+1} \quad . \quad (39)$$

The probability that a new call (which is not blocked) on a g-type platform requires a hand-off is therefore

$$b(g) = \sum_{i=1}^{N(g)} \rho_n(g,i) \cdot b(g,i) \quad . \quad (40)$$

All calls that are successfully handed off renew service (in the target cell) in the first dwell time phase. So the probability that a call that has been handed off requires yet another hand-off is $b(g,1)$.

The probability that the call is forced to terminate on its k^{th} hand-off attempt is

$$Y(g,k) = b(g) \cdot P_H(g) \cdot \{ b(g,1) \cdot [1-P_H(g)] \}^{k-1} . \quad (41)$$

The forced termination probability is therefore

$$P_{FT}(g) = \sum_{k=1}^{\infty} Y(g,k) . \quad (42)$$

This can be compactly written in closed form as

$$P_{FT}(g) = b(g) \cdot P_H(g) / [1 - \psi(g)] , \quad (43)$$

in which we have let

$$\psi(g) = b(g,1) \cdot [1-P_H(g)] . \quad (44)$$

In [13] it is shown that the forced termination probability can also be written in the convenient form

$$P_{FT}(g) = \frac{[\bar{T}(g) / \bar{T}_D(g)]}{1 + \Gamma^{-1}(g) \cdot P_H(g)} \quad (45)$$

in which

$$\Gamma(g) = \{ [\pi(g)]^{-N(g)} - 1 \} . \quad (46)$$

In addition it is shown that tight bounds on $P_{FT}(g)$ are given by

$$\frac{u}{1 + [u / S(1/u)]} \geq P_{FT} \geq \frac{u}{1 + u} . \quad (47)$$

in which by definition

$$S(z) = (e^z - 1) / z , \quad (48)$$

and

$$u = [\bar{T}(g) / \bar{T}_D(g)] \cdot P_H(g) . \quad (49)$$

For convenience we denote the bracketed term in (49) by δ . Thus, $u = \delta \cdot P_H$. The lower bound on P_{FT} is exact for the case of ned dwell time, and the upper bound is especially tight for (constant) deterministic dwell time. The quantity, u , however, depends on $P_H(g)$, and must be found using the computational procedure described previously. Thus, for $u \leq 10^{-3}$, P_{FT} is given (correctly to 3 significant figures) by $\delta \cdot P_H$. Similarly for $u \geq 10^3$, P_{FT} is given (correctly to 3 significant figures) by $\delta \cdot P_H / (1 + \delta \cdot P_H)$.

Hand-Off Activity Factor

We define the hand-off activity factor, $\eta(g)$ as the expected number of hand-off attempts for a non-blocked call on a g -type platform. Let $U(g,k)$ be the probability that a g -type call requires exactly k hand-off attempts before ending either by (satisfactory) completion or by forced termination. Exactly one hand-off is needed if a new call completes the necessary number of dwell time phases before call completion and then either fails on its first hand-off attempt or succeeds on this attempt but is satisfactorily completed before completing the $N(g)$ additional dwell time phases for the next hand-off attempt. Recall that a new call begins in dwell time phase i with probability $\rho_n(g,i)$ and that $N(g)-i+1$ phases would have to be completed to generate the first hand-off attempt for a new call that begins in the i^{th} dwell time phase. Also recall that if the first hand-off is satisfactory, service of the call is renewed (in the target cell) in the first dwell time phase.

Let $\varphi(g)$ be the probability that a g -type call which requires a hand-off will NOT require an ADDITIONAL hand-off. Thus

$$\varphi(g) = P_H(g) + [1-P_H(g)] \cdot [1-b(g,1)] , \quad (50)$$

which is algebraically identical to $1 - \psi(g)$.
Then we find

$$\begin{aligned} U(g,1) &= b(g) \cdot \varphi(g) \\ U(g,2) &= b(g) \cdot [1-P_H(g)] \cdot b(g,1) \cdot \varphi(g) , \\ U(g,3) &= b(g) \cdot \{ [1-P_H(g)] \cdot b(g,1) \}^2 \cdot \varphi(g) , \end{aligned} \quad (51)$$

and, in like manner

$$U(g,k) = b(g) \cdot \varphi(g) \cdot [\psi(g)]^{k-1} . \quad (52)$$

The hand-off activity factor is

$$\eta(g) = \sum_{k=1}^{\infty} k \cdot U(g,k) . \quad (53)$$

This can be compactly written in closed form as

$$\eta(g) = b(g) \cdot \varphi(g) / [1-\psi(g)]^2 = b(g) / [1-\psi(g)] . \quad (54)$$

Note that this is the same factor that appears in equation (43). Some bounds on $\eta(g)$ are derived in [13].

DETECTION OF HAND-OFF NEEDS: EXTENDED PERFORMANCE MEASURES

As mentioned previously, an important aspect of the hand-off process is the detection of the need of a communicating platform for a hand-off. Up to this point, this issue has been set aside - only the limitations imposed by the availability of communication resources in the target cell have been considered. Many alternative methods for *hand-off need detection* can be devised [14]-[17]. Consideration of such methods is not the subject of this paper.

Here it is assumed that *some* algorithm is used to initiate the hand-off procedure. For the present purpose, the algorithm is characterized by a parameter, ζ , which denotes the probability that a need for a hand-off is *missed*. We consider a hand-off need to be *missed* if: the conditions necessitating a platform hand-off are either not detected; or, are not detected early enough to permit the necessary exchange of supervisory signals before the current link fails. Of course, there is also a non-zero probability that a hand-off request is initiated when in fact a hand-off is NOT needed. Such requests cause unnecessary system *churning* but do not result in forced terminations. This is because even if the request cannot be accommodated in the target cell, the call can be continued in the platform's current cell. Therefore *false requests* are not considered here. We suggest that hand-off initiation schemes consider the trade-off between *missed hand-off needs* and *system churning* and characterize the former by a probability, ζ , as described above.

Because of *missed hand-off needs*, only the fraction, $1-\zeta$, of actual hand-off departures will result in the initiation of hand-off attempts. Thus for a *homogeneous system*, the average hand-off attempt and hand-off departure rates (per cell) will not necessarily be equal - rather they will be related by $\Lambda_h = (1-\zeta) \cdot \Delta_h$. For the more general (homogeneous or) non-homogeneous case the relationship is

$$\Lambda_h = \theta \cdot (1-\zeta) \cdot \Delta_h \quad . \quad (55)$$

A *need* for hand-off will result in a forced termination if either the detection of that need is missed, or the need is detected but resources to service it are unavailable in the target cell. Thus the probability that a needed hand-off results in a forced termination is

$$P_H^*(g) = \zeta + (1-\zeta) \cdot P_H(g) \quad . \quad (56)$$

The expressions for overall performance measures for the more general case are similar to those described in the previous sections except that $P_H(g)$ in equations (46)-(54) should be replaced by $P_H^*(g)$ as given by (56).

Specifically, given a *hand-off need detection scheme* for which ζ is known, the calculation for the more general case proceeds as follows. The function $Q_2(\Lambda_h)$ should be taken as

$$Q_2(\Lambda_h) = \Lambda_h - \theta \cdot (1-\zeta) \cdot \Delta_h \quad , \quad (57)$$

and the algorithm outlined in the section on *COMPUTATIONAL PROCEDURE* should be followed to find the statistical equilibrium solution.

The resulting state probabilities are then used to determine the carried traffic components and blocking probabilities using (32)-(35). The hand-off need failure probabilities, $P_H^*(g)$, are calculated using (37) and (56). The forced termination probabilities are calculated using

$$P_{FT}^*(g) = b(g) \cdot P_H^*(g) / [1 - \psi^*(g)] \quad , \quad (58)$$

in which

$$\psi^*(g) = b(g,1) \cdot [1 - P_H^*(g)] \quad . \quad (59)$$

In this case, the hand-off activity factor for a g-type call is the expected

number of NEEDED hand-off attempts generated by a non-blocked call on a g-type platform. This is given by

$$\eta^*(g) = b(g) / [1 - \psi^*(g)] \quad . \quad (60)$$

Overall traffic performance characteristics that account for the reliability of the hand-off initiation algorithm as well as the limitations due to communications resource availability and organization were calculated using these formulas.

DISCUSSION OF RESULTS

Performance Characteristics

Numerical results were generated using the approach described in this paper. The homogeneous case ($\theta=1$) with *perfect* hand-off need detection ($\zeta=0$) was used for Figures CS:2.X and 3.X. Figures CS:2.X are for C=36 channels and two platform types (G=2). Figures CS:3.X are for C=15 channels and a single user type (G=1). Figures CS:4.X are for the homogeneous case with *imperfect* hand-off need detection, C=15 channels and two platform types (G=2). Additional figures for a situation with C=25 channels and three platform types (G=3); high mobility, low mobility, and stationary are given in [13]; and for C=15 and G=2, in [19].

For all figures an unencumbered call duration of 100 sec. was assumed and only a *channel limit*, C, (no quotas), was considered. The figures presented are of two kinds. One kind has an abscissa reflecting call demand, with dwell times held fixed. In these the abscissa is new call origination rate for platform type 1 (denoted $\Lambda(1)$) with the ratio of new call origination rates from other platform types held fixed with respect to type 1. The other kind has an abscissa reflecting platform dwell times (with new call origination rates held fixed) and with dwell times of other platform types held in fixed ratio to that of type 1. Because of the scaling in these figures, the abscissa can also be envisioned as proportional to cell size.

For Figures CS:2.X, and 4.X, two platform types (G=2) were considered, a low mobility platform type and a high mobility platform type. The mean dwell times of these platform types were taken in the ratio of 5 to 1, throughout. For an appropriate cell size, these choices can represent pedestrians and autos, respectively. Figures CS:3.X, which are for a single platform type, show the effects of different dwell time distributions (all having the same mean but with different coefficients of variation). For convenient reference, the specific parameters used for each figure are included in a List of Figures.

Figure CS:2.1 shows blocking and forced termination probabilities as a function of new call arrival rate from type 1 platforms. Since there are no channel quotas, the blocking probability is the same for each platform type. However since the platforms have different (mobility parameters) mean dwell times, there are differences in forced termination probability. Increasing the value of C_h , reduces forced termination probability at the cost of increasing blocking probability. It is seen that this allows a very favorable exchange for the purpose of reducing forced termination probabilities. As C_h increases, forced termination probabilities decrease dramatically while blocking probability increases relatively slowly.

Figure CS:2.2 shows the total carried traffic per cell and the traffic components from each platform type. For low demand, the carried traffic increases linearly with increasing demand. For higher demand the increase in traffic is less than the proportional increase in demand. This is especially true for large C_h since blocking performance is sacrificed to accommodate hand-offs. From Figure CS:2.4, it is seen that this decrease in carried traffic depends on the ratio $\delta = \bar{T}(g) / \bar{T}_D(g)$. As C_h increases, there are compensating trends. For values of $\delta \gg 1$, there tends to be many hand-off attempts. So the trend toward a decrease in carried traffic due to increased C_h (and increased blocking) is offset by a reduction in forced termination probability, which allows more calls to be carried to successful completion.

Figures CS:2.3, 2.4, and 2.5 show dependence of performance on dwell time for a type 1 platform with the ratio of dwell times held fixed at 5 to 1. If one considers dwell time to be proportional to cell size, these figures can be interpreted as showing dependence of the performance measures on cell size (but with demand in a cell being held fixed). In these figures the value of $\Lambda(1) = 2.7E-04$ calls/sec. was assumed. (Roughly this corresponds to a level of traffic in which every user makes one call during the busy hour.) In Figure CS:2.3 we see that for small values of the abscissa, blocking probability increases with increasing dwell time, and from Figure CS:2.4 carried traffic increases as well. For forced termination probabilities, there are two opposite effects. Increasing the dwell time tends to reduce hand-off attempts and therefore to reduce forced termination probability. But then more calls are sustained to completion so the cells carry more traffic, and this tends to increase forced termination probability. The net effect is an increase in the forced termination probability for small dwell times (cell size). With further increasing cell size, the carried traffic tends to saturate, but hand-off activity continues to decrease. Thus forced termination probability decreases.

Figure CS:2.4 shows the carried traffic per cell. It is seen that for small cell size, ($\delta \gg 1$) most of the total carried traffic is from the less mobile platform type. This is because the high mobility platforms (since they require more hand-offs) are more likely to encounter a hand-off failure. For large cell sizes ($\delta \ll 1$), there is in the limit no impact due to different mobility characteristics, because all calls will be completed in the cell where they originated. Thus (with equal new call origination rates) the platform types are equally likely to get and retain a channel. The carried traffic of each platform type tends to the same value.

Figure CS:2.5 shows hand-off activity of each platform type. For the parameter choices shown, this was found to be essentially independent of C_h , but strongly dependent on cell size. This is because in the usual range of interest, $P_H \ll 1$, so essentially all hand-off attempts succeed and the average number of hand-off attempts is determined mainly by mobility parameters.

In Figures CS:3.1 and 3.2 the effects of different dwell time pdf's are considered. There is only one platform type. For a given abscissa in Figure CS:3.1, the mean dwell time is held fixed and blocking and forced termination probabilities are shown for values of $N(1)=1,2,3,4$. These values correspond to squared coefficients of variation of 1, 0.5, 0.333, and 0.250, respectively. Increasing $N(1)$ corresponds to dwell times tending more toward the deterministic case. It is seen that both blocking and forced termination probabilities tend to increase as dwell time becomes less random. The effect of different numbers of phases is most noticeable for small cell size. Unless

the dwell time is somewhat less than the unencumbered session holding time, the predominant parameter that affects the performance characteristics is mean holding time. It is seen that performance is relatively insensitive to the dispersion of the dwell time (about its mean). *This fortuitous result allows one often to calculate relevant performance characteristics using the negative exponential model for dwell times. This leads to computations requiring the smallest number of states within this framework.*

Figure CS:3.2 shows hand-off activity for $C_h=0,2,4$, and $N(1)=1,2,3$. The mean dwell time was taken to be 20 sec. At low demand the hand-off activity is close to $100/20 = 5$. This is the ratio of unencumbered session duration to mean dwell time and roughly the expected number of cell boundary crossings for a call that is sustained to completion. This agrees with the mathematical limits developed in [13]. As demand increases, the hand-off activity tends to decrease because hand-off failure becomes more likely, so calls tend to be terminated before many crossings can be made. It is also seen that as C_h increases, hand-off activity is increased because calls are less likely to be terminated early.

In Figures CS:4.1 the effects of imperfect detection of hand-off need is considered. The missed hand-off initiation probability, ζ , was taken as $1.0E-03$. We found blocking probability to be insensitive to this parameter over the interesting range of demand. However, for forced termination probability, it is seen that in the low demand region, the missed detection probability determines the forced termination probability. This is because (in this region) few calls are terminated due to lack of resources in the target cell. For high demand the lack of resources dominates the forced termination probability so the curves tend to follow those for perfect detection. Thus the missed detection probability sets a "floor" on forced termination performance. Significant improvement in P_{FT} by increasing C_h is only attainable in the high demand region. Additional figures and comparisons are given in [13].

Extrapolation to Larger Systems

Because the number of states needed to compute performance characteristics can be unmanageable for large C , it is useful to consider how theoretical performance characteristics can be generated for such cases. This can be done by normalizing the new call offered traffic per cell to the number of channels per cell and proceeding with the calculations as described in previous sections for several manageable values of C . For this purpose the new call offered traffic is taken as the product of the *new call origination rate per cell* and *unencumbered holding time*. Blocking and forced termination probabilities (for various values of C_h) are plotted against C using a semi-logarithmic grid. This is shown in Figure CS:5 for a fixed new call origination rate of 0.7 erl/channel per cell. It can be seen that the performance curves are essentially straight lines when displayed on this plot. This shows that for the chosen parameters, blocking and forced termination probabilities decrease exponentially with increasing C . In Figure CS:5, points for values of C less than or equal to 30 were computed using a stand-alone desktop Sun workstation. Points for C between 35 and 40 were computed using remote access to the Cornell Supercomputer Facility. The curves for C greater than 40 were obtained by simple linear extrapolation. Higher order polynomial extrapolations are possible but seem unnecessary for the purpose of theoretical prediction of performance characteristics. Thus it can be seen that computing several points in the range of C from 20 to 30, allows simple and excellent extrapolation of performance characteristics to predict performance of larger

systems, and obviates a major computational effort.

CONCLUSIONS

The framework that we are developing using multi-dimensional birth-death processes can be used to compute theoretical traffic performance characteristics for cellular communication systems in which mobile platforms of different types are simultaneously present. A broad class of dwell time distributions can be accommodated. With proper interpretation, the essential features of the radio environment that have an impact on traffic performance can be incorporated in the dwell time variable. Traffic performance characteristics are dependent on the mean of the dwell time that a mobile platform spends within communications range of a wireless gateway, but are relatively insensitive to the variance.

The effects of imperfect hand-off need detection on overall traffic performance is considered in an extended framework. This allows major issues of the hand-off problem to be studied separately - one aspect dealing with how the hand-off procedure is initiated and the other dealing with communications traffic and spectral resource allocations. The extended framework allows the computation of combined overall traffic performance characteristics for various alternative schemes for each of these issues.

A method is described whereby traffic performance characteristics can be simply extrapolated to large systems, so intensive computations are generally not necessary.

Cut-off priority for hand-off calls in cellular systems allows a very favorable exchange between blocking and forced termination performance. Thus, it can be effectively used in this context.

ACKNOWLEDGEMENTS

The author is grateful to Mr. Gustavo MONTE for skillful computer programming that generated the performance characteristics shown.

The research reported in this paper was supported in part by the U.S. National Science Foundation under Grant No. NCR-9025131 and in part by IST/SDIO under Grant No. N00014-91-J-4063 administered by the U.S. Office of Naval Research. Some of the results were obtained using the Cornell National Supercomputer Facility, a resource of the Center for Theory and Simulations in Science and Engineering (Cornell Theory Center), which receives major funding from the U.S. National Science Foundation and IBM Corporation, with additional support from New York State and members of the Corporate Research Institute.

REFERENCES

- [1] D.C. Cox and D.O. Reudink, "Layout and control of high-capacity systems," Chap. 7 in *Microwave Mobile Communications*, ed. by W.C. Jakes, Jr., Wiley:New York, 1974, pp. 545-622.
- [2] D. Hong and S.S. Rappaport, "Traffic model and performance analysis for cellular mobile radiotelephone systems with prioritized and non-prioritized hand-off procedures," *IEEE Trans. Vehic. Technol.*, vol. VT-35, pp. 77-92, Aug. 1986.
- [3] S.A. El-Dolil, W-C. Wong, and R. Steele, "Teletraffic performance of highway microcells with overlay macrocell," *IEEE J. Select. Areas Commun.*, vol. 7, no. 1, pp. 71-78, Jan. 1989.
- [4] D. Hong and S.S. Rappaport, "Priority oriented channel access for cellular systems serving vehicular and portable radio telephones," *IEE (British) Proc., Part I, Commun., Speech and Vision*, vol. 136, pt. I, no. 5, pp. 339-346, Oct. 1989.
- [5] R. Guerin, "Queueing-blocking system with two arrival streams and guard channels," *IEEE Trans. Commun.*, vol. COM-36, pp. 153-163, Feb. 1988.
- [6] S.S. Rappaport, "The multiple-call hand-off problem in high-capacity cellular communications systems," *IEEE Trans. Vehic. Technol.*, vol. 40, no. 3, pp. 546-557, Aug. 1991.
- [7] S. Rappaport, "Models for call hand-off schemes in cellular communication networks," in *Third Generation Wireless Information Networks*, ed. by S. Nanda and D.J. Goodman, Kluwer Academic Publishers: Boston, 1992, pp. 163-185.
- [8] S.S. Rappaport, "Modeling the hand-off problem in personal communications networks," *Proc. IEEE Vehicular Technology Conference, VTC '91*, St. Louis, May 1991, pp. 517-523.
- [9] R.B. Cooper, *Introduction to Queueing Theory, 2nd ed.*, Elsevier North Holland:New York, 1981.
- [10] R. Steele, "Deploying personal communication networks," *IEEE Communications Magazine*, September 1990, pp. 12 ff.
- [11] D.C. Cox, "Personal communications - a viewpoint," *IEEE Communications Magazine*, November 1990, pp. 8 ff.
- [12] A.J. Viterbi, "Wireless digital communication: a view based on three lessons learned," *IEEE Communications Magazine*, September 1991, pp. 33 ff.
- [13] S.S. Rappaport, "Blocking, hand-off and traffic performance for cellular communication systems with mixed platforms and imperfect hand-off initiation," CEAS Technical Report no. 660, College of Engineering and Applied Sciences, State University of New York, Stony Brook, New York 11794-2350, November 26, 1991, Appendix.
- [14] A. Murase, I.C. Symington, and E. Green, "Handover criterion for macro and microcellular systems," *Proc. IEEE Vehicular Technology Conference, VTC '91*, St. Louis, May 1991, pp. 524-530.
- [15] S.T.S. Chia, "The control of handover initiation in microcells," *Proc. IEEE Vehicular Technology Conference, VTC '91*, St. Louis, May 1991, pp. 531-536.
- [16] M. Gudmundson, "Analysis of handover algorithm," *Proc. IEEE Vehicular Technology Conference, VTC '91*, St. Louis, May 1991, pp. 537-542.
- [17] O. Grimlund and B. Gudmundson, "Handoff strategies in microcellular systems," *Proc. IEEE Vehicular Technology Conference, VTC '91*, St. Louis, May 1991, pp. 505-510.
- [18] F.P. Kelly, *Reversibility and Stochastic Networks*, Wiley:New York, 1979.
- [19] S.S. Rappaport, "Communications Traffic Performance for Cellular systems with Mixed Platform Types," *Workshop Record, Winlab Workshop on Third*

Generation Wireless Information Networks," East Brunswick, NJ, April 28-29, 1992, pp. 131-158. Available from Rutgers University, WINLAB, Box 909, Piscataway, NJ 08855-0909, USA.

APPENDIX A: COMPUTATIONAL PROCEDURE

An important consideration is that the quantities, F'_g , Λ_h , and Δ_h which are reflected in the determination of the $q(i,j)$'s in the flow balance equations depend on the unknown state probabilities, $p(s)$. Therefore the flow balance equations are actually a set of simultaneous nonlinear equations. However, by beginning with guesses for Λ_h and F'_g , ($g=1,2,3,\dots,G$), the iterative approach of [6] can again be used. In this approach the system is linear within each iteration. To explain the procedure we define two functions, Q_1 and Q_2 .

Function $Q_1(\Lambda_h)$

The function $Q_1(\Lambda_h)$ is a vector function which, for given Λ_h , returns; all state probabilities, $p(s)$, $s=0,1,2,3,\dots,s_{\max}$; fractions of g -type call hand-off attempts, F'_g ; and the average hand-off departure rate, Δ_h . The function is defined by an algorithmic calculation as follows:

Step 1: Given Λ_h , begin with guesses or previous values for F'_g . Solve the flow balance equations, for the state probabilities.

Step 2: Using the solution determine the new fractions F'_g and compare with the previous values. If the relative error in any of the $p(s)$'s or fractions, F'_g , exceeds the requirement (say 10^{-4} for 3 significant figures), repeat step 1 using the latest values for F'_g . Otherwise, for the given Λ_h and the solution, determine the resulting average hand-off departure rate, Δ_h , and return the latest values.

For the numerical computations presented here, a modified Gauss-Seidel algorithm was used to solve the flow balance equations in step 1. This allows using only one flow balance equation at a time and tends to reduce the number of coefficients that have to be stored. This advantage is at the expense of having to do additional calculations to regenerate the coefficients for each equation as needed. The algorithm is especially useful if the the number of states is very large. Other methods for solving simultaneous linear equations can be used. In step 2, the method of Successive Substitutions was used to find the F'_g 's as described above.

Function $Q_2(\Lambda_h)$

The function $Q_2(\Lambda_h)$ is a scalar function which, for given Λ_h , calls the function, $Q_1(\Lambda_h)$, and returns the difference, $\Lambda_h - \Delta_h$. That is,

$$Q_2(\Lambda_h) = \Lambda_h - \Delta_h \quad . \quad (A.1)$$

The calculation proceeds by using some algorithm to find the root of $Q_2(\Lambda_h) = 0$. For this value of Λ_h the hand-off arrival and departure rates are equal as required (in the homogeneous case).

For non-homogeneous cellular systems, the same basic approach can be used except that the arrival and departure rates are, in general, unequal. Instead, they are related by some (non-unity) constant that specifies the "tilt" or degree of inhomogeneity for the cell under consideration. That is,

we require

$$\Lambda_h = \theta \cdot \Delta_h \quad , \quad (A.2)$$

in which, $\theta \neq 1$, is a given "tilting parameter." The parameter is chosen to be greater or less than unity to characterize a cell which on the average experiences more or less hand-off arrivals than departures, respectively. One can proceed as described above, except that the function Q_2 should be defined as

$$Q_2(\Lambda_h) = \Lambda_h - \theta \cdot \Delta_h \quad . \quad (A.3)$$

Any of a number of root finding algorithms can be used to solve the equation $Q_2(\Lambda_h)=0$, to the desired accuracy. For the numerical computations presented here, the Bisection Method was used.

It is important to note that if an iteration scheme such as Gauss-Seidel is used in step 1, the overall procedure requires three levels of iterations. One in finding the root of Q_2 , another in finding the F_g 's in Q_1 , and a third in solving for the state probabilities in step 1. Furthermore, because solving for the state probabilities is part of the innermost iterative loop, fairly intensive computations can be required to produce results. Nevertheless we were able to generate numerical performance characteristics for many non-trivial parameter choices using readily available workstations. For some additional parameter choices, calculations were done via remote access to the Cornell National Supercomputing Facility.

Some examples of cpu time requirements needed to obtain an equilibrium solution are given in Table 2. Computation of the first point in a run is more time consuming because the computer program must first identify and order the states and their predecessors. This part of the calculation need not be repeated for subsequent points. The required cpu time for a point also depends on particular parameter values. These effects produce the spread shown in the table.

TABLE 2: APPROXIMATE REQUIRED COMPUTATION TIME
G=2, N(1)=2, N(2)=2, CHANNEL LIMIT = C.

C	number of states	CPU TIME IN MINUTES			
		Sun SparcStation 1st pt.	SLC other pts.	Supercomputer 1st pt.	other pts.
15	3,876	1.32 - 6.73	1.03 - 5.34	0.43 - 2.26	0.26 - 1.72
25	23,751	-	-	7.96 - 26.65	2.33 - 16.61

APPENDIX B: DERIVATION OF SOME FORMULAS

For ease of notation we temporarily omit showing explicit dependence on the platform type in the formulas below. That is we take $\pi \equiv \pi(g)$, $N \equiv N(g)$, etc.

A key ratio that arises in the analytical development is $b(g) / [1-\psi(g)]$. To find $b(g)$ we substitute (38) and (39) into (40) and recall that $\rho_n(g,i) = 1/N(g)$. The resulting geometric progression can be summed to give

$$b(g) = \frac{\pi (1-\pi^N)}{N (1-\pi)} \quad . \quad (B.1)$$

Then using (38)-(40) in (44) gives

$$1-\psi(g) = 1-\pi^N + \pi^N \cdot P_H \quad . \quad (B.2)$$

Thus

$$\frac{b(g)}{1-\psi(g)} = \frac{\pi}{N (1-\pi)} \frac{(1-\pi^N)}{(1-\pi^N) + \pi^N \cdot P_H} \quad . \quad (B.3)$$

Dividing numerator and denominator by $1-\pi^N$ gives

$$\frac{b(g)}{1-\psi(g)} = \frac{\pi}{N (1-\pi)} \frac{1}{1 + \frac{\pi^N \cdot P_H}{(1-\pi^N)}} \quad . \quad (B.4)$$

The denominator of (B.4) can be rewritten. Then using (38) we find

$$\frac{b(g)}{1-\psi(g)} = \frac{\mu_D}{N\mu} \frac{1}{1 + [P_H / (\pi^{-N} - 1)]} \quad . \quad (B.5)$$

Similarly from (38) we get

$$1 / (\pi^{-N} - 1) = \mu_D^N / [(\mu_D + \mu)^N - \mu_D^N] \quad . \quad (B.6)$$

Expand the denominator of (B.6) using the binomial formula and subtract the leading term. Then divide numerator and denominator of (B.6) by μ^N . This results in

$$1 / (\pi^{-N} - 1) = (\mu_D / \mu)^N / \sum_{j=1}^N {}^N C_j \cdot (\mu_D / \mu)^{N-j} \quad , \quad (B.7)$$

in which ${}^N C_j$ denotes the combinatorial $\{N! / [j! (n-j)!]\}$. Now we define the parameter δ , as the ratio of mean unencumbered session time to mean dwell time. That is,

$S(1/u) \geq S(\theta)$. So a looser upper bound in (B.16) yields

$$\frac{\delta}{1 + u/S(1/u)} \geq \eta(g) \geq \frac{\delta}{1 + \delta \cdot P_H} . \quad (\text{B.17})$$

It is noted from (43) that $\eta(g) = b(g) / [1 - \psi(g)]$. It follows from (43) and (B.17) that the forced termination probabilities are bounded by

$$\frac{u}{1 + u/S(1/u)} \geq P_{FT}(g) \geq \frac{u}{1 + u} . \quad (\text{B.18})$$

The upper bound in (B.18) stems from letting N get very large and therefore corresponds to the deterministic case for which it is especially tight. The upper bound in (B.16) is exact for this case. The lower bound in (B.18) is exact for $N=1$ which corresponds to ned dwell time.

Using these bounds, it can be shown that for $u \leq 10^{-3}$, P_{FT} is given (correctly to 3 significant figures) by $\delta \cdot P_H$. Similarly for $u \geq 10^3$, P_{FT} is given (correctly to 3 significant figures) by $\delta \cdot P_H / (1 + \delta \cdot P_H)$. Furthermore the upper and lower bounds approach the same limit $u/(1+u)$ as $u \rightarrow \infty$ and as $u \rightarrow 0$. The ratio of the difference between the upper and lower bounds to the lower bound reaches a maximum value for a certain value of u say u^* where $u^* = 1/w^*$, and w^* satisfies the implicit equation

$$w^* = \log(1 + 2 \cdot w^*) . \quad (\text{B.19})$$

This can be used to find a *uniform upper bound* (independent of N and u) on the percent error in using the bounds of (B.18). However the bounds given for the ranges above are probably more useful in practical cases.

APPENDIX C: RELATIONSHIP BETWEEN DWELL TIMES, CELL SIZES AND PLATFORM TYPES

For the purpose of determining approximate relationships between dwell times, cell sizes and platform types we consider regular hexagonal cells having a radius, R . The area of such a cell is

$$\text{area} = (3\sqrt{3} / 2) R^2 \quad . \quad (\text{C.1})$$

Define an "equivalent " circle having the same area. The radius of the equivalent circle is

$$R_{\text{eq}} = (3\sqrt{3} / 2\pi)^{1/2} R \quad . \quad (\text{C.2})$$

Now assume that the path of a platform in a cell is a straight line. Thus the path traversed traces a randomly placed chord of the equivalent circle. It is not difficult to show that the *mean* length of such a chord is

$$\bar{L} = (4/\pi) R_{\text{eq}} \quad . \quad (\text{C.3})$$

If the platform moves at a constant speed V , the dwell time is

$$\bar{T}_D = \bar{L} / V \quad . \quad (\text{C.4})$$

Substituting (C.2), and (C.3) into (C.4) so as to eliminate \bar{L} , yields the following relationship between cell radius, speed and dwell time:

$$R = (4/\pi) (3\sqrt{3} / 2\pi)^{-1/2} \bar{T}_D \cdot V \quad . \quad (\text{C.5})$$

The relationship is plotted in Figure CS:1.1.

LIST OF FIGURES

Figure CS:1.1: Dwell times, cell sizes and platform types.

Figure CS:1.2: A terrestrial cellular communication system with mixed platform types.

Figure CS:2.1: Blocking and forced termination probabilities depend on demand.

$$C=36, C_h=0,3,6, G=2, N(1)=2, N(2)=2, v(1,0)=v(2,0)=350$$

$$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$$

$$\bar{T}_D(1,1)=\bar{T}_D(1,2)=500 \text{ sec.}, \bar{T}_D(1)=1000 \text{ sec.},$$

$$\bar{T}_D(2,1)=\bar{T}_D(2,2)=100 \text{ sec.}, \bar{T}_D(2)=200 \text{ sec.}$$

Figure CS:2.2: Traffic carried depends on demand.

$$C=36, C_h=0,3,6, G=2, N(1)=2, N(2)=2, v(1,0)=v(2,0)=350$$

$$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0,$$

$$\bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$$

$$\bar{T}_D(1,1)=\bar{T}_D(1,2)=500 \text{ sec.}, \bar{T}_D(1)=1000 \text{ sec.},$$

$$\bar{T}_D(2,1)=\bar{T}_D(2,2)=100 \text{ sec.}, \bar{T}_D(2)=200 \text{ sec.}$$

Figure CS:2.3: Blocking and forced termination probabilities depend on dwell time means.

$$C=36, C_\theta=0,3,6, G=2, N(1)=2, N(2)=2, v(1,0)=v(2,0)=350,$$

$$\Lambda(1)=\Lambda(2)=2.70E-04$$

$$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$$

$$\bar{T}_D(1,1)=\bar{T}_D(1,2), \bar{T}_D(2,1)=\bar{T}_D(2,2),$$

$$\bar{T}_D(2) / \bar{T}_D(1) = 0.2 .$$

Figure CS:2.4: Traffic carried depends on dwell time means.

$$C=36, C_\theta=0,3,6, G=2, N(1)=2, N(2)=2, v(1,0)=v(2,0)=350,$$

$$\Lambda(1)=\Lambda(2)=2.70E-04$$

$$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$$

$$\bar{T}_D(1,1)=\bar{T}_D(1,2), \bar{T}_D(2,1)=\bar{T}_D(2,2),$$

$$\bar{T}_D(2) / \bar{T}_D(1) = 0.2 .$$

Figure CS:2.5: Hand-off activity depends on dwell time means.

$$C=36, C_\theta=0,3,6, G=2, N(1)=2, N(2)=2, v(1,0)=v(2,0)=350,$$

$$\Lambda(1)=\Lambda(2)=2.70E-04$$

$$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$$

$$\bar{T}_D(1,1)=\bar{T}_D(1,2), \bar{T}_D(2,1)=\bar{T}_D(2,2),$$

$$\bar{T}_D(2) / \bar{T}_D(1) = 0.2 .$$

Figure CS:3.1: Blocking and forced termination probabilities

depend on dwell time means and variances.

$C=15, C_h=0,2,4, G=1, N(1)=1,2,3,4, v(1,0)=300,$

$\Lambda(1)=2.75E-04, \bar{T}(1)=100 \text{ sec.}, \bar{T}_D(1,i)=\bar{T}_D(1,j)$

Figure CS:3.2: Hand-off activity for different dwell time pdf's

depends on demand.

$C=15, C_h=0,2,4, G=1, N(1)=1,2,3, v(1,0)=300,$

$\Lambda(1)=2.75E-04, \bar{T}(1)=100 \text{ sec.}, \bar{T}_D(1,i)=\bar{T}_D(1,j),$

Figure CS:4.1: Effect of imperfect detection of hand-off need on blocking

and forced termination probabilities. (poor detection, $\zeta=1.0E-03$).

$C=15, C_h=0,2, G=2, N(1)=2, N(2)=2, v(1,0)=150,$

$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$

$\bar{T}_D(1,1)=\bar{T}_D(1,2)=500 \text{ sec.}, \bar{T}_D(1)=1000 \text{ sec.},$

$\bar{T}_D(2,1)=\bar{T}_D(2,2)=100 \text{ sec.}, \bar{T}_D(2)=200 \text{ sec.}$

Figure CS:5 Blocking and forced termination probabilities depend on channel limit.

$G=2, N(1)=2, N(2)=2, C_h=0,2,4,6. \text{ new calls} = 0.7 \text{ erl/channel}$

$\alpha_n(2)=\Lambda_n(2)/\Lambda_n(1) = 1.0, \bar{T}(1)=\bar{T}(2)=100 \text{ sec.},$

$\bar{T}_D(1,1)=\bar{T}_D(1,2)=500 \text{ sec.}, \bar{T}_D(1)=1000 \text{ sec.},$

$\bar{T}_D(2,1)=\bar{T}_D(2,2)=100 \text{ sec.}, \bar{T}_D(2)=200 \text{ sec.}$

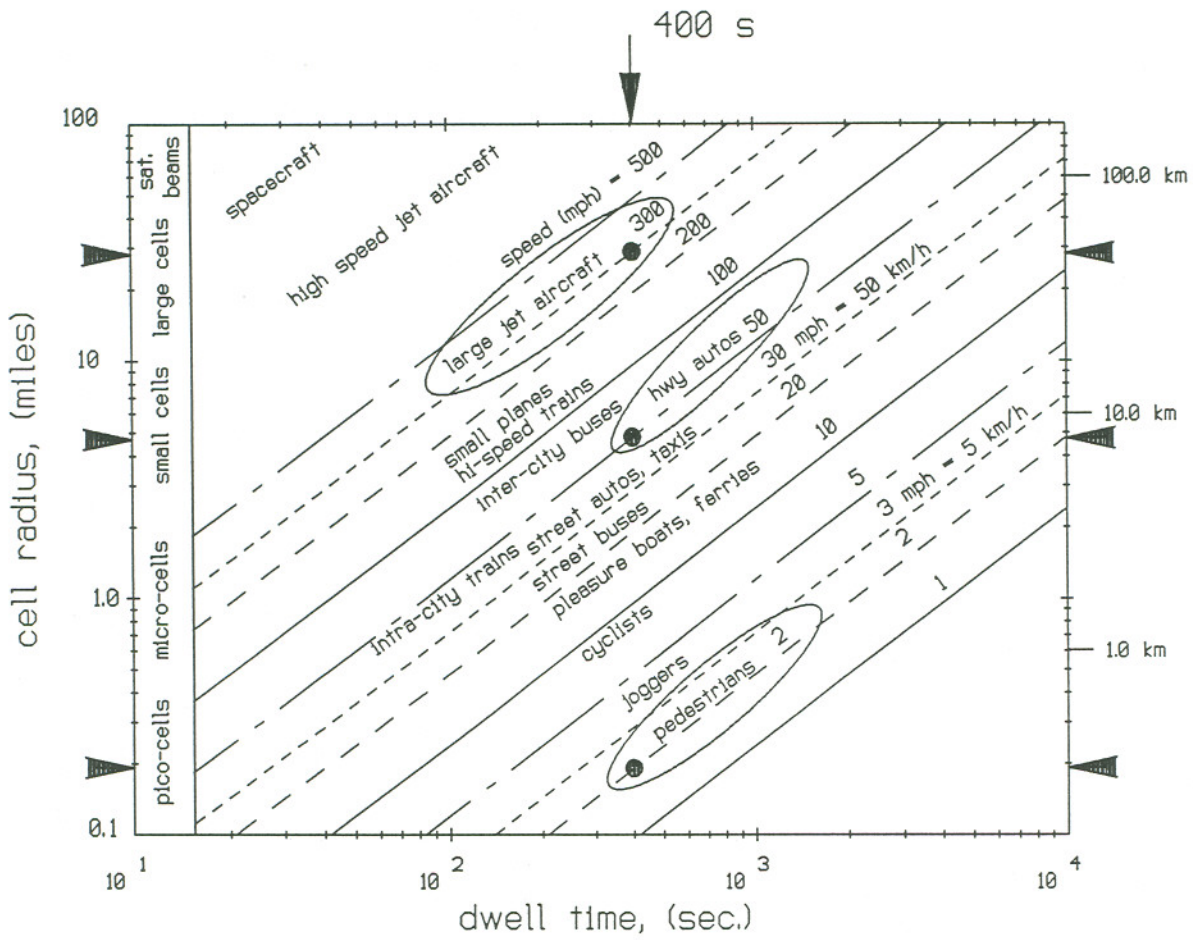
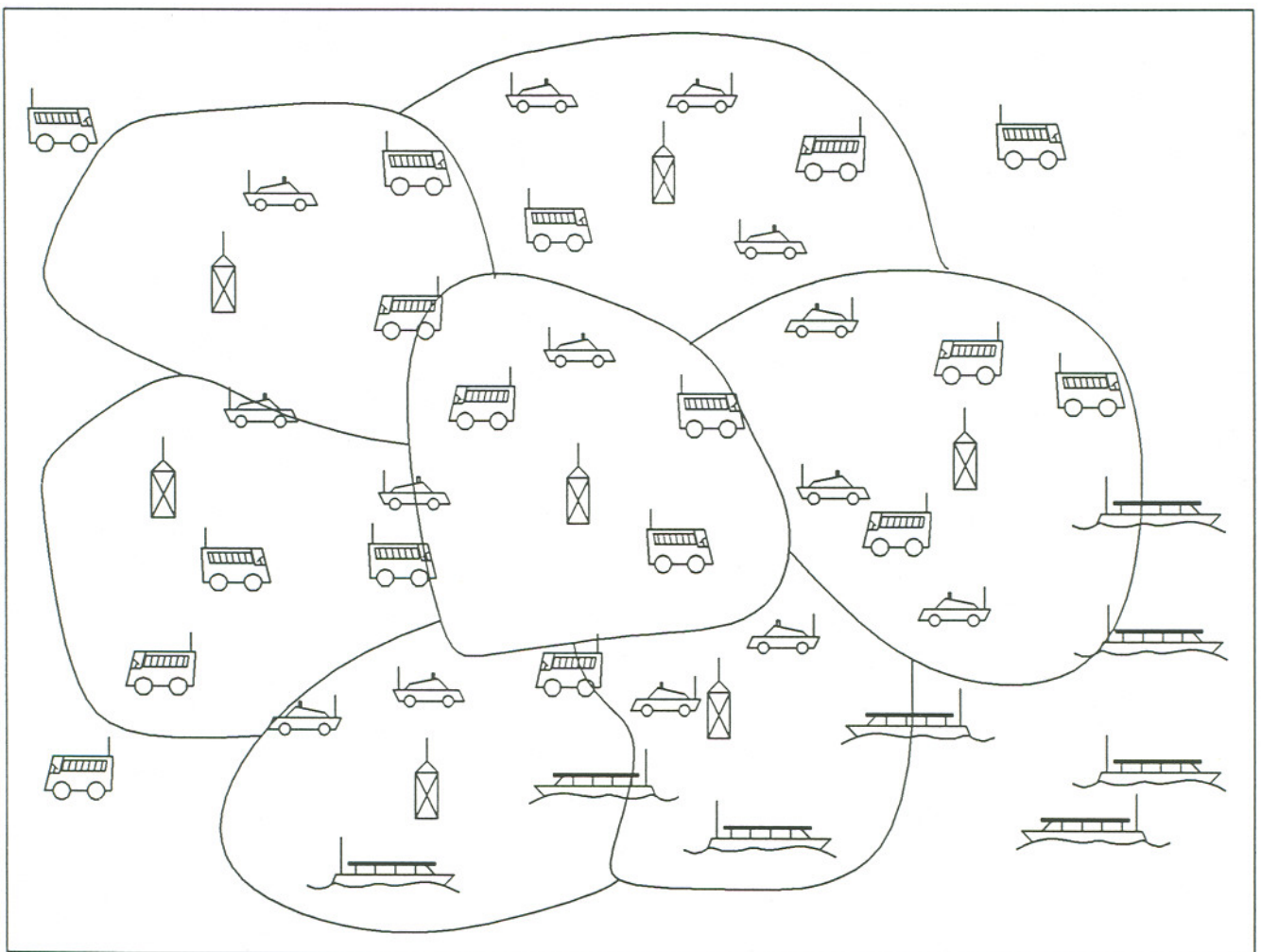


Figure CS:1.1 Dwell Times, Cell Sizes, and Platform Types

Figure CS:1.2

A terrestrial cellular communication system with several platform types.



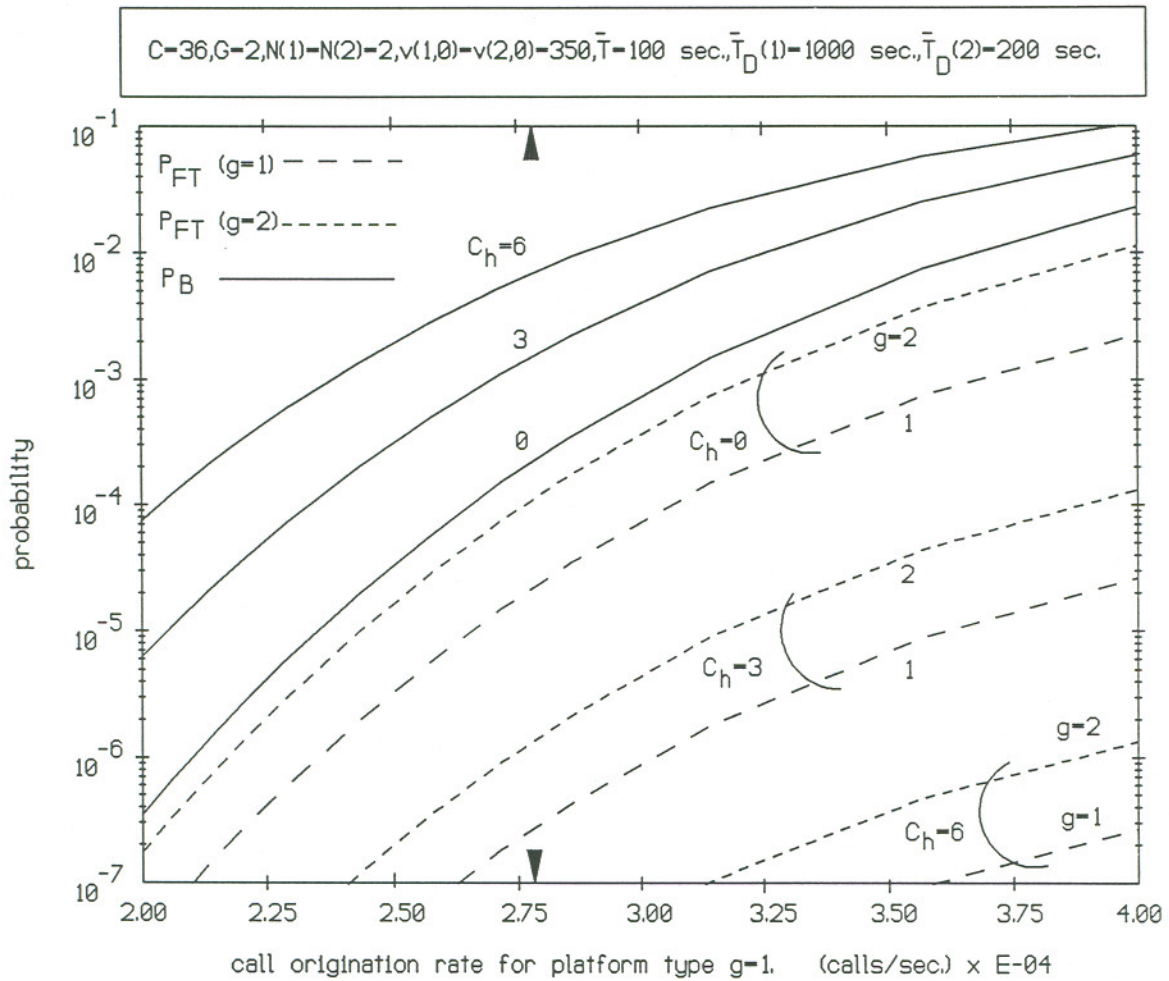


Figure CS:2.1 Blocking and Forced Termination Probabilities Depend on Call Demand

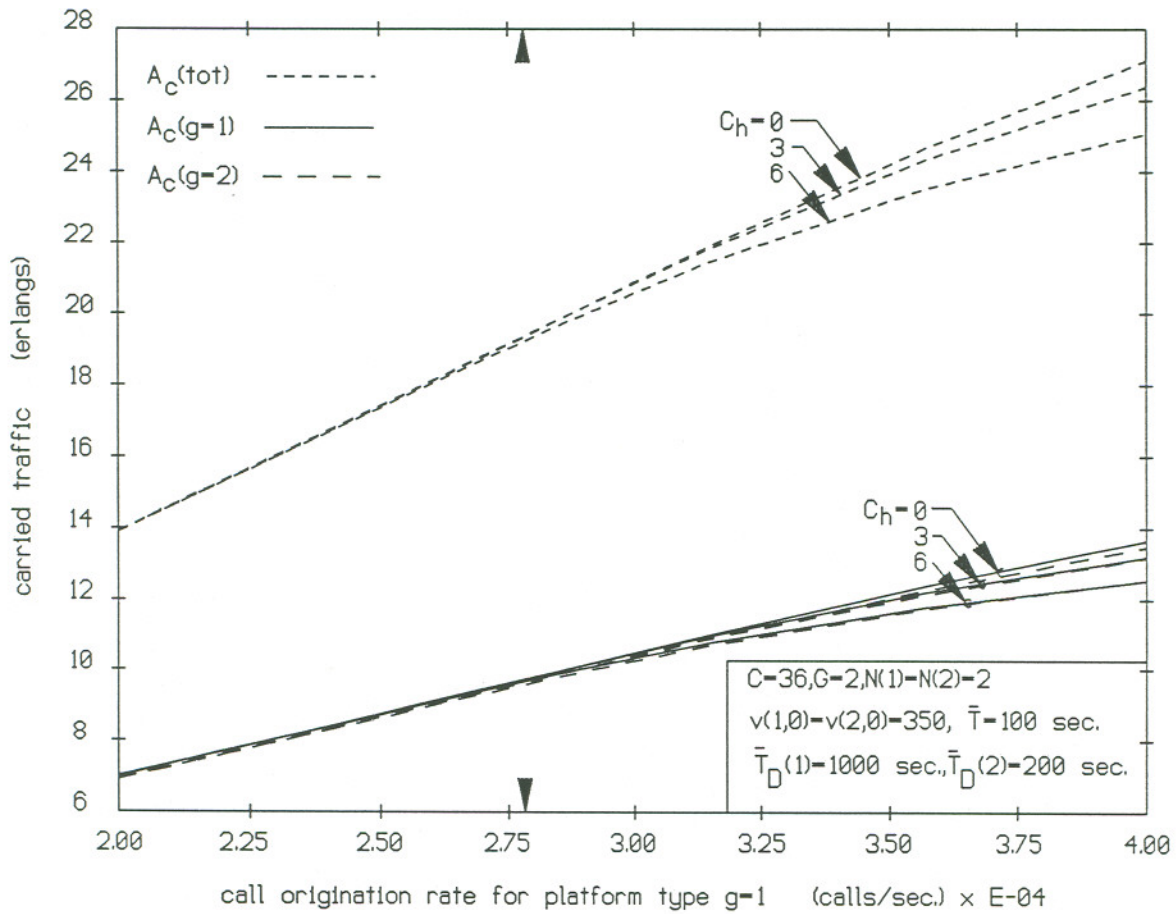


Figure CS:2.2 Traffic Carried Depends on Call Demand

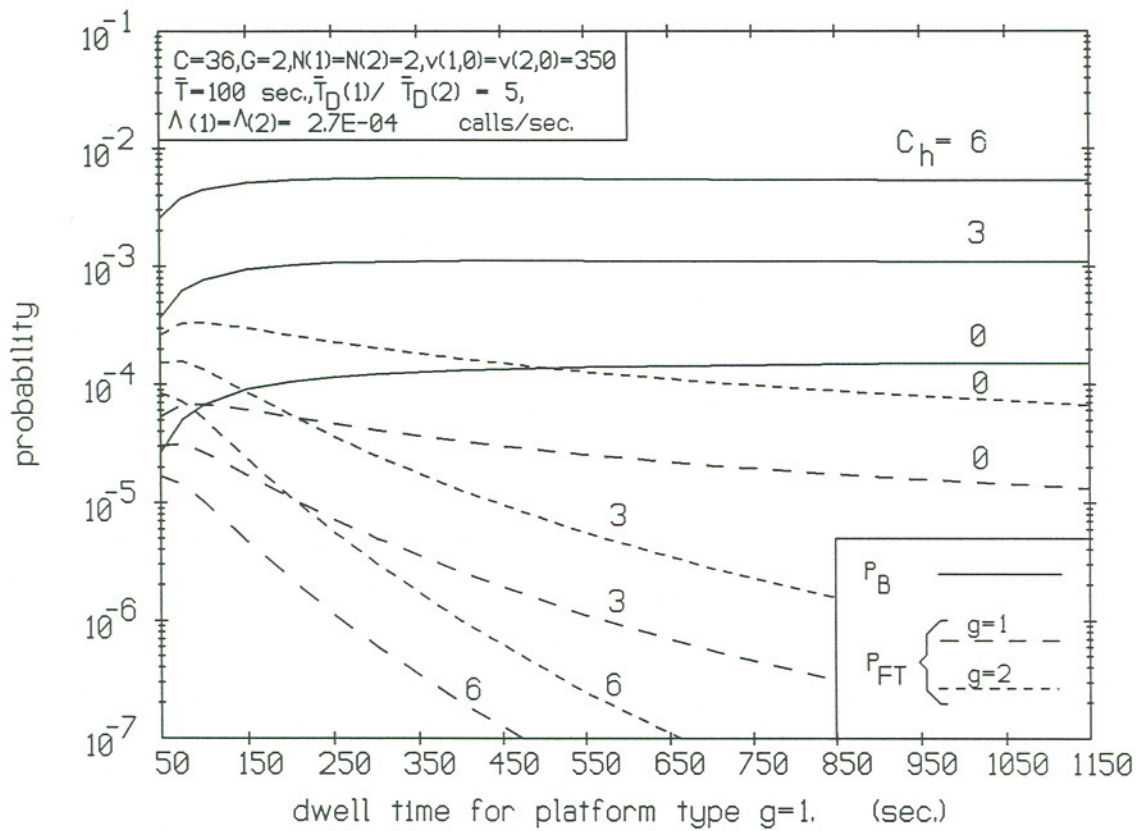


Figure CS:2.3 Blocking and Forced Termination Probabilities Depend on Dwell Time Means

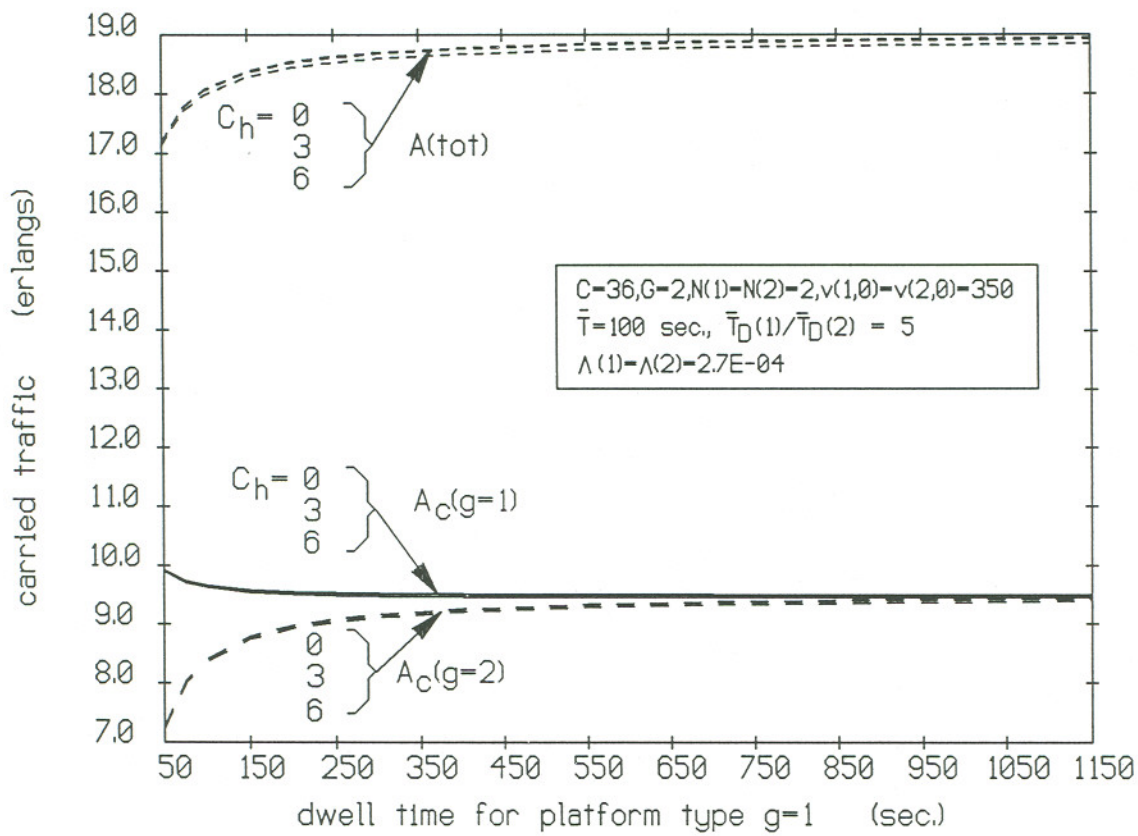


Figure CS:2.4 Traffic Carried Depends on Dwell Time Means

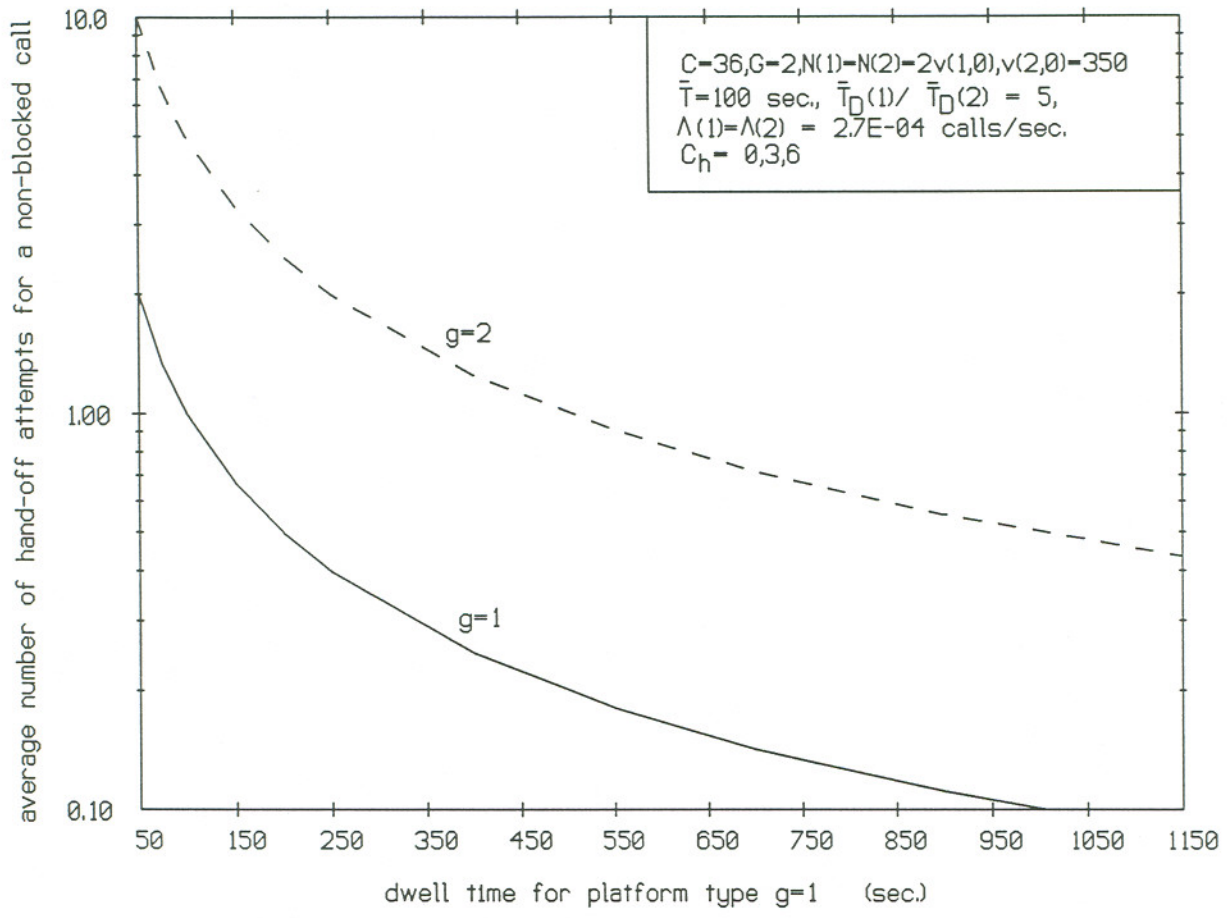


Figure CS:2.5 Hand-Off Activity Depends on Dwell Time Means

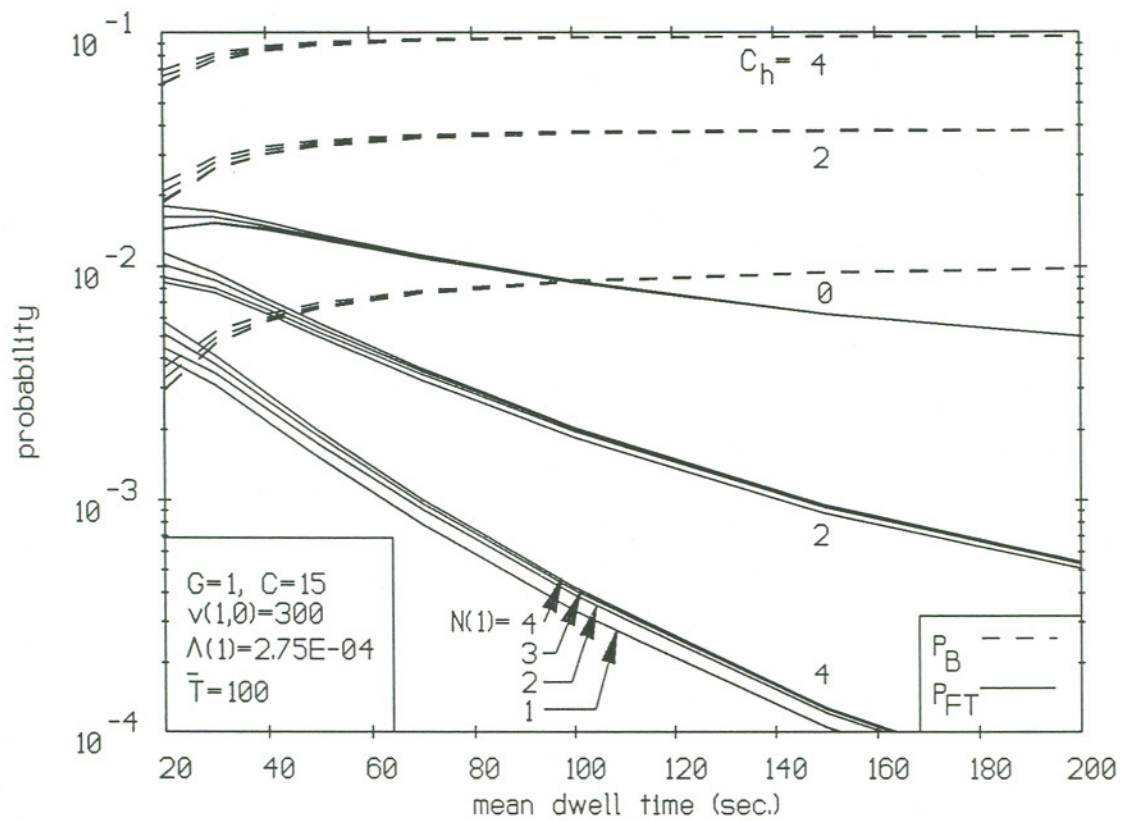


Figure CS:3.1 Blocking and Forced Termination Probabilities Depend on Dwell Time Means and Variances

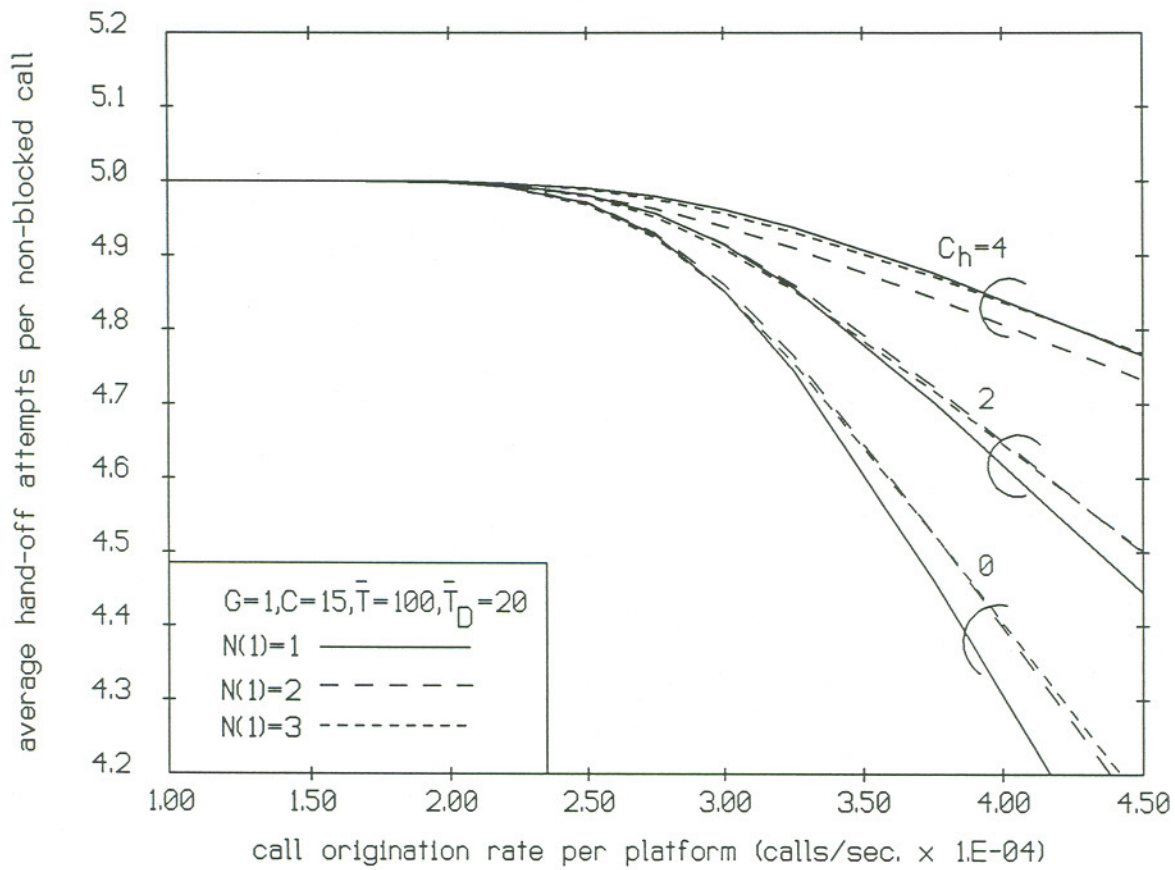


Figure CS:3.2 Hand-Off Activity for Various Dwell Time PDF'S

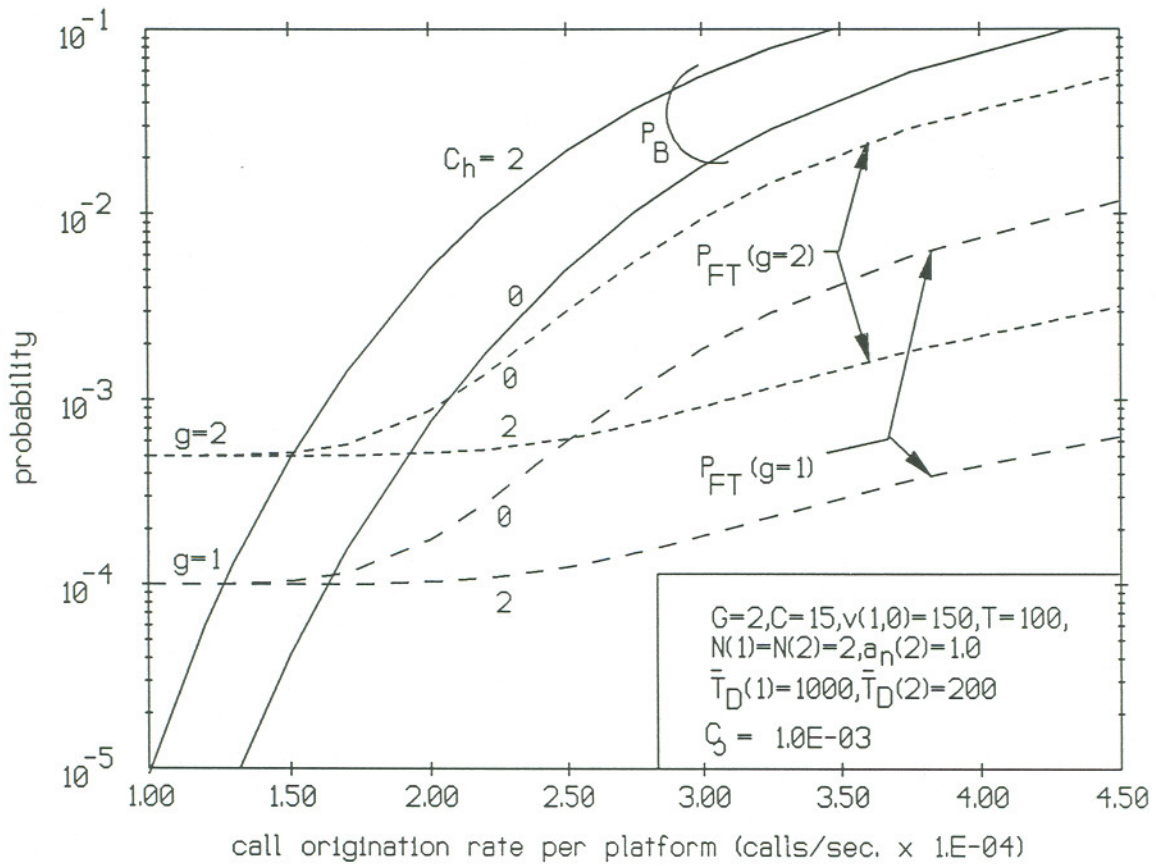


Figure CS:4.1 Effect of Imperfect Hand-Off Initiation on Performance

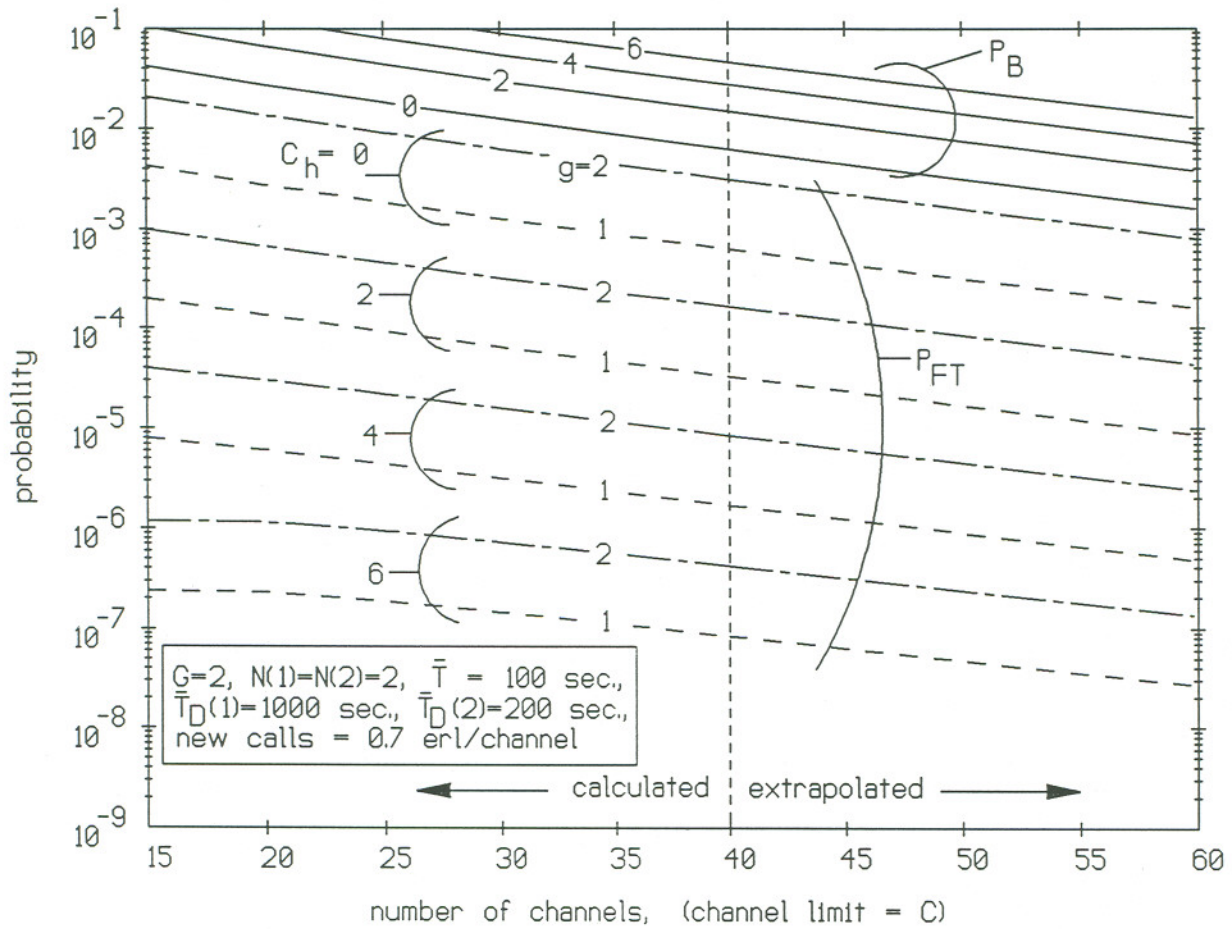


Figure CS:5 Blocking and Forced Termination Probabilities Depend on Channel Limit