State University of New York at Stony Brook
College of Engineering and Applied Sciences

Technical Report No. 760

Session Oriented Data Communications for Mobile Computing in Cellular Systems with Multiple Traffic Classes

by

Yunsang Park and Stephen S. Rappaport

Department of Electrical and Computer Engineering
State University of New York
Stony Brook, New York 11794-2350

e-mail: park@sbee.sunysb.edu, rappaport@sunysb.edu

Date: May 29, 1998

Session Oriented Data Communications for Mobile Computing in Cellular Systems with Multiple Traffic Classes

Yunsang Park and Stephen S. Rappaport

Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794-2350

park@sbee.sunysb.edu, rappaport@sunysb.edu

Keywords : Cellular communications, mobile computing, traffic performance analysis.

ABSTRACT

A scheme is proposed for cellular communication systems in which voice and data co-exist. Each traffic class is managed using a session-oriented strategy. During a session a mobile user has access to network resources, although this access may be shared with others. A session is typified as a either a voice session or a data session. Preemptive priority is used to guarantee transparency for voice sessions. For data sessions, the scheme attempts to maintain a connection to the network. Data sessions, which are preempted or which are disconnected doing the hand-off process, are allowed a fixed number of reconnection attempts. Only after the given number of such attempts to reconnect have failed, is the data session deemed to have failed. The strategy attempts to maintain connectivity for mobile data users in a way that is transparent to them. This allows for example, *mobile computing* users to continue functioning autonomously (though not indefinitely) in an off-line mode. We develop a tractable analytical model for traffic performance based on multidimensional birth-death processes. This approach allows consideration of mixed traffic types as well as mixed platform types, (such as pedestrians, automobiles, and buses), which may have very different mobility characteristics. Theoretical traffic performance characteristics are calculated. These include carried traffic, blocking probability, forced termination probability, average time per suspension, and the average number of suspensions per session.

The research reported in this paper was supported in part by the U.S. National Science Foundation under Grant no. NCR 94-15530 and in part by BMDO/IST under grant no. N00014-9511217 administered by the U.S. Office of Naval Research. General research support from Hughes Network Systems is gratefully acknowledged.

INTRODUCTION

Originally developed to provide voice service, cellular communication technology now seeks to accommodate additional services. With the rapid development of computer industry and the explosively increasing demand for computer communication, the need to accommodate data services is of continually increasing importance. The limited availability of wireless resources and the hostile environment of wireless channel, necessitate special consideration of problems that do not arise in fixed wired networks.

Hand-off is the process of changing the access point of a mobile for the purpose of maintaining quality of radio link as the mobile moves. Unsuccessful hand-off causes premature termination of a call. Because of the potential for premature termination, the wireless segment can be a bottle-neck for ATM or other data services that are accommodated on the backbone [1], [2]. The following issues arise: 1) How to minimize the likelihood of premature termination when diverse traffic types coexist in a system? 2) How to guarantee transparency for time sensitive calls even if a variety calls share limited wireless resources? 3) What strategy should be used for managing heterogeneous call traffic types in order to guarantee Quality of Service (QoS) for all traffic classes?

The characterization of hand-off performance as an important system metric was considered in [3], which also treated prioritized schemes for hand-off calls. More general user mobility and mixed platform types were considered in [4]. Mixed call types and mixed platform types were considered in [5].

A scheme to send data traffic on cellular networks is considered in [6]. In that scheme, data type calls use channels only during idle times, which are defined as the time intervals between consecutive voice calls. In this scheme, transparency of data traffic is guaranteed for voice call users. Also, in [7], some strategies for management of data calls in this "background service" are considered. The analysis of traffic performance of this system is studied in [8], where more complicated issues which arise from the existence of different mobility classes are studied. Prioritized hand-off procedures were also considered.

Considering the time insensitive nature of many data traffic types, some delay during communication is not critical factor for acceptable QoS. However, a lost connection which results in the termination of a session is significant, since it waste valuable wireless resources. Therefore, the question of "how to maintain connectivity of a mobile user to the network?" is an important issue. Systems that attempt to maintain mobile user connectivity to the network by automatically and transparently attempting to reconnect disrupted links are considered in [9]. A disconnection need not necessarily result in a failed session with discarded information and wasted resource usage. Only a single traffic class, namely time insensitive data, was treated.

In the present paper, a system that uses widely disparate call traffic types is considered. A session-oriented approach is used. A session is typified as either a voice session or a data session. Preemptive priority is used to guarantee transparency for voice sessions. For data sessions, the scheme attempts to maintain a connection to the network. Data calls, which are preempted or which are disconnected doing the hand-off process, are allowed a fixed number of reconnection attempts. Only after the given number of such attempts to reconnect have failed, is the data session deemed to have failed. The strategy attempts to maintain connectivity for mobile data users in a way that is transparent to them. This allows for example, *mobile computing* users to continue functioning autonomously (though not indefinitely) in an off-line mode. We examine the combined effect of mixed call traffic types in the same system and quantify the performances of each type. We consider mixed platform types. That is platforms with different mobility characteristics are present in the system. In addition, cut-off priority is used to reserve some resources for voice sessions, since they are especially susceptible to hand-off failures.

The model for traffic performance can be cast in the analytic framework that has been developing in recent years [3], [4], [8], [9]. The approach, which uses multidimensional birth-death processes, allows numerical computation of relevant state probabilities [4]. These probabilities are used to compute important traffic performance measures for the proposed system.

SESSION MANAGEMENT STRATEGY

A voice session characterized by real time conversation typically needs a continuous uninterrupted session. A data session is generated by a source that carries and processes data type information. Often data transfers are time insensitive. Voice mail does not need real time transmission, on the other hand is transmitted by a data service and. We consider this to be a data session.

When the physical link between a mobile terminal that has an active data session and network fails, the data session is *suspended* and mobile terminal will attempt to reconnect by successive reconnection requests made at random time intervals. These are called ***reconnection attempts***. A maximum number of reconnection attempts, N, is allowed for each suspended session. If a reconnection has not been secured after this maximum is reached, the session is considered to have failed and is cleared from the system. The number of reconnection attempts for suspended sessions is counted and updated in a counter in the mobile terminal.

It is assumed that there is a maximum number, H, of suspended sessions that the system will allow in each cell. It is possible that a platform with a suspended session on board leaves its current cell at this epoch a reconnection attempt is made to establish a link in the new call. This hand-off attempt counts towards the limit, N. If, in the target cell, there are no available channels to accommodate the arriving data session, and if there are also already H suspended sessions in the target cell, the arriving data session cannot be admitted in the target cell. So, even if a suspended session has not exhausted the allowable number of reconnection attempts, it will be forced into termination if it fails its hand-off attempts and the system already has H suspended session in the target cell. A detailed description of mobile terminal's reconnection counter is given in the APPENDIX *A*.

Since voice sessions must be transmitted or received on a real time basis, reconnection attempts are not allowed. Instead, voice sessions have preemptive priority over data sessions for using channel resources. When a voice session arrives and finds all channels occupied, an active data session (if any are present) will be either suspended or terminated to accommodate it. The choice of which data session to be suspended or be terminated is assumed to be random. If there are no active data sessions that can be preempted to service the incoming voice session the voice arrival will not be accommodated. That is, it will be blocked if it is a new call, or terminated if it is a hand-off.

There are two possible reasons that cause an active data session to be suspended. One is failure of a hand-off attempt of the data session. Specifically, if a data session attempts a hand-off when the channels in the target cell are fully occupied but, in the target cell, there are fewer than H data sessions suspended, the hand-off attempt will be fail but the session will wait for another connection opportunity as a suspended session. The other reason for suspension arises when an active data session is preempted by an arriving voice session. When a voice session arrives in a cell in which all channels are occupied and fewer than H sessions are in suspension, and, at least, one active session is of data type, an arriving voice session will obtain a connection but an active data session will be suspended.

In the following we let g be an index that defines the platform type and mobility. Consider a suspended session that has already failed k-1 reconnection attempts. The next attempt is called the "k-reconnection attempt" where $1 \le k \le N$. It is important to emphasize that there are two driving processes that generate ***reconnection attempts***. One is the ***retry*** process, which consists of successive statistically independent realizations of a random variable, $T_r(k,g)$, to generate epochs for retry attempt times for a suspended session. The other is the ***hand-off*** departure process -- because *hand-off attempts always try to establish a link and therefore count as reconnection attempts*. The random variable, $T_r(k,g)$, gives the time from the previous ***reconnection event*** (either hand-off or retry) to the next anticipated ***retry attempt***. The random variable, $T_r(k,g)$, can in general depend on k. Thus, the minimum rate of reconnection attempts depends on the number of attempts that have already been made. Of course, if the supporting platform leaves its current cell before the anticipated retry epoch, a hand-off attempt (to establish a link) will be made at that time and the value of k will be adjusted. If the session is in a suspended state after this attempt, a new random variable (for a retry epoch) will be generated. The random variable, $T_r(k,g)$, generated after the k-1 reconnection attempt, represents the maximum time to the next anticipated retry attempt. This is called the "*k*-trial time". The next reconnection attempt will be made either at this time, or at the time that the supporting platform leaves the cell, whichever is the shortest. A suspended session that has not reestablished a link after *k-1* reconnection trials and is waiting for the next (k^{th}) reconnection attempt, is called a "*k*-suspended session".

There are three possible reasons that cause a data session to be prematurely terminated. One is failure of a hand-off attempt for the data session. If a hand-off for a data session is attempted to a cell in which all channels are fully occupied and H suspended sessions exist, it will fail and the session, whether it is active or suspended, must be terminated. The second reason is that an active data session is preempted by an incoming voice session. When a voice session arrives in a cell in which all channels are fully occupied, H suspended sessions exist, and, at least, one session is of data type, the arriving voice type session will obtain a connection but an active data session will be terminated. The third reason that the allowable number of reconnection attempts for the suspended data session has been exhausted. A suspended data session is allowed up to N reconnection attempts. N consecutive failures of reconnection attempts will cause termination of the data session.

MODEL DESCRIPTION

In the following description, we freely borrow concepts and phraseology put forth in [3] and [4]. It is suggested that readers review those papers.

We consider a large cellular system with many mobile platforms of several types. Each mobile can potentially generate a voice session or a data session. However, each platform can support at most one connection at any give time and each connection needs one channel (resource) to communicate. The platform types differ primarily in there mobility characteristics. The maximum number of simultaneous connections that each base station can support is C .

When a platform with either an active or suspended session moves to another cell, a hand-off is needed. We assume hand-off detection and initiation are perfect. For a voice session, a hand-off attempt will succeed to gain a connection in the target cell if there are fewer than C voice sessions in that cell. A voice session that fails to gain a connection will lose its wireless link and cleared from system. For a data session, a hand-off attempt will gain access to a connection in the target cell if there are less than C sessions, either of voice type or data type, in progress in that cell. When a hand-off of a data session fails to acquire a connection, a session will be suspended if the reconnection counter in the terminal indicates less than N attempts *and* no more than H suspended sessions are in the same cell.

Platform mobility is characterized using the concept of *dwell time* [4], [8] - a random variable which is defined as the duration of time that a two-way communication link of satisfactory quality can be maintained between a platform and its current base, for whatever reason. The dwell time of platform in a cell depends on many factors including; mobility, signal power, propagation conditions, fading, etc. [4]. The amount of time that a session must use a channel for satisfying communication is modeled using the concept of *unencumbered session duration*. The unencumbered session duration is a random variable, which is the amount of time that the call would spend in service if there were no suspensions or forced termination. The unencumbered data session duration depends on the amount of data and speed of modem. Similarly, the k-trial time is a random variable. A k-suspended session will execute a retry attempt after the epoch of the k-trial time unless it moves to another cell. If a k-suspended session moves to another cell before the epoch of the k-trial time, a hand-off attempt will be made.

We emphasize here that the term ***reconnection attempt*** applies to any attempt of a *suspended session* to reestablish a link to the network. Reconnection attempts are initiated by attempted hand-off of suspended session as well as by the *retry* process of suspended session.

EXAMPLE PROBLEM STATEMENT

The system can support G type mobile platforms, indexed by $\{g = 1, 2, 3......., G\}$. No more than one session can be supported by a platform at any given time. The voice session origination rate from a noncommunicating g-type platform is denoted $\Lambda_v(g)$. We define $\alpha(g) = \Lambda_v(g) / \Lambda_v(1)$. The data session generation rate from a noncommunicating g-type platform is denoted $\Lambda_w(g)$. It is, also, defined $\beta(g) = \Lambda_w(g) / \Lambda_v(g)$. Potentially, a platform can generate two types of sessions, voice and data. The number of noncommunicating g-type platforms in any cell is denoted by $v(g,0)$. Therefore, the total voice session generation rate for g-type platforms in any cell can be denoted $\Lambda_{nv}(g) = \Lambda_v(g) \times v(g,0)$ and the total data session generation rate for g-type platforms in a cell can be denoted $\Lambda_{nw}(g) = \Lambda_w(g) \times v(g,0)$. It is assumed that the number of noncommunicating platforms is much larger than the number of connections that a cell can support so that the session origination rate does not depend on the number of calls being served (This is called an infinite population model) [4].

Generally the bandwidth and other resources that needed for connection of a data session may vary each session. A model that considers resource use based on connection type is developed in [5]. It is assumed that each active connection, either for a data session or a voice session, requires the same amount of resources. Each cell or gateway can support maximum of C connections. There are no quotas for either specific mobility platform type or specific type of session. Cut-off priority for hand-off activity (either of voice sessions or data sessions) and reconnection attempts for data sessions are included in the present discussion. Thus, C_h connections in each cell are reserved for hand-off attempts (for either voice sessions or data sessions) and for reconnection attempts (only for suspended data sessions) in the cell. A connection will be established for an arriving voice session *only* if there are less than $C\text{-}C_h$ *active voice sessions* in the cell. For an arriving data session, a connection will be made if there are fewer than C sessions, either of voice type or data type, in the cell. Hand-off attempt of a voice session will fail if there are C *active voice sessions* in the cell. The voice session that fails its hand-off attempt will be terminated and cleared from the system. A hand-off attempt of a data type session will fail if there are C sessions, *either of voice type or data type*, in the cell. The data session that fail its hand-off attempt will be suspended if its reconnection counter indicates less than $N+1$ *and* less than H suspended sessions are in the cell.

The platform is considered to "leave" the cell at the expiration of its current (random) dwell time. A communicating platform that leaves a cell generates a hand-off arrival to some other cell. The dwell time in a cell for a g-type platform is a ned random variable, $T_D(g)$, having a mean $\overline{T}_D(g) = 1/\mu_D(g)$. This dwell time distribution can be generalized [10], [11], [12]. The unencumbered voice session duration on a g-type platform is a ned random variable, $T_v(g)$, having a mean $\overline{T}_v(g) = 1/\mu_v(g)$. The unencumbered data session duration on a g-type platform is a ned random variable, $T_w(g)$, having a mean $\overline{T}_w(g) = 1/\mu_w(g)$. The k-trial time of a suspended session on g-type platform is ned random variable, $T_r(g)$, having a mean $\overline{T}_r(k,g) = 1/\mu_r(k,g)$, where $1 \le k \le N$, and $\mu_r(k,g)$ (k=1, 2, ...,N;g=1,.....,G) is the parameter that determines the reconnection attempt rate for a *k-suspended session* on a g-type platform.

STATE DESCRIPTION

Considering a single cell, we define the ***cell state*** by a sequence of non-negative integers. When a maximum of N reconnection attempts are permitted for a suspended data session, the state of the cell can be written as G n-tuples as follows

$$\begin{matrix} v_1 & w_1 & r_{1,1} & r_{1,2} & r_{1,3} & \cdots & r_{1,N} \\ v_2 & w_2 & r_{2,1} & r_{2,2} & r_{2,3} & \cdots & r_{2,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_G & w_G & r_{G,1} & r_{G,2} & r_{G,3} & \cdots & r_{G,N} \end{matrix} \qquad (1)$$

where v_g {g=1,2,......,G} is the number of active voice sessions on g-type platforms, w_g {g=1,2,.......,G} is the number of active data sessions on g-type platforms, and $r_{g,k}$ {g=1,2,.....,G;k=1,2,.....,N} is the number of k-suspended sessions on g-type platforms. For convenience, we order the states using an index s = 0, 1,......, $S_{\max}$. Thereafter, v_g, w_g, and $r_{g,k}$ can be written explicitly dependent on the state. That is $v_g = v(s,g)$, $w_g = w(s,g)$, and $r_{g,k} = r(s,g,k)$.

When the cell is in state s, the following characteristics can be determined. The number of voice sessions is

$$v(s) = \sum_{g=1}^{G} v(s, g). \tag{2}$$

The number of active data sessions is

$$w(s) = \sum_{g=1}^{G} w(s, g). \tag{3}$$

The number of suspended sessions on g-type platforms is

$$r(s, g) = \sum_{k=1}^{N} r(s, g, k). \tag{4}$$

The number of suspended sessions, regardless of platform type, is

$$r(s) = \sum_{g=1}^{G} r(s, g). \tag{5}$$

The total number of data sessions on g-type platforms in a cell, either active or suspended, is

$$a(s, g) = w(s, g) + r(s, g). \tag{6}$$

The total number of data sessions, regardless of platform type, either active or suspended, is

$$a(s) = \sum_{g=1}^{G} a(s, g). \tag{7}$$

The number of active sessions, regardless of type of session, is

$$J(s, g) = v(s, g) + w(s, g). \tag{8}$$

And, the total number of sessions in progress in a cell is

$$J(s) = v(s) + w(s). \tag{9}$$

There are constraints on permissible cell states. These includes the total number of active sessions in a cell must be fewer than or equal to maximum supportable connections, $J(s) \le C$; and the total number of suspended sessions in a cell must be fewer than or equal to the maximum number of suspended sessions allowed in a cell, $r(s) \le H$.

Given that the system is in state s, the conditional probability, $Q_s(s, g)$, that an active data session held on *a g-type platform* will be suspended when an voice session arrives is given by

$$Q_s(s,g) = \begin{cases} w(s,g)/w(s) & \text{if } w(s) \neq 0,\ r(s) < H \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

The conditional probability, $Q_t(s,g)$, that an active data session held on *a g-type platform* will be chosen for termination when an voice session arrives is given by

$$Q_t(s,g) = \begin{cases} w(s,g)/w(s) & \text{if } w(s) \neq 0,\ r(s) = H \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

DRIVING PROCESSES

There are nine possible driving processes. We use Markovian assumptions for driving processes to allow solution within the multidimensional birth-death process framework [3], [4], [9], [11]. Each process is listed below

- $\{nv\}$: generation of voice sessions
- $\{nw\}$: generation of data sessions
- $\{cv\}$: completion of voice sessions
- $\{cw\}$: completion of data sessions
- $\{hv\}$: hand-off arrival of voice sessions
- $\{hw\}$: hand-off arrival of data sessions
  - $\{hw_0\}$: active data session hand-off arrivals
  - $\{hw_1\}$: 1-suspended session hand-off arrivals
  - $\vdots$ $\vdots$
  - $\{hw_N\}$: N-suspended session hand-off arrival
- $\{dv\}$: hand-off departures of voice sessions
- $\{dw\}$: hand-off departures of data sessions
  - $\{dw_0\}$: active data session hand-off departures
  - $\{dw_1\}$: 1-suspended session hand-off departures
  - $\vdots$ $\vdots$
  - $\{dw_N\}$: N-suspended session hand-off departures
- $\{r\}$: retry attempt

The dimensions of session generation or session completion processes, either of voice type or data type, are G, respectively, since there are G different types of mobility platforms. Similarly, the dimensions of the hand-off arrival and departure processes, either of voice type or data type, are $G\times(N+1)$, since N- times of reconnection attempts are permitted for a suspended session besides of existence of active sessions. The dimension of the retry process is $G\times N$, since N types of suspended data sessions can be arise from G different types of platforms.

Compete discussion of each driving process is given in [13].

FLOW BALANCE EQUATIONS

The total transition flow into state s from any permissible predecessor state x can be written as

$$\begin{aligned} q(s,x) &= \gamma_{nv}(s,x)+\gamma_{nw}(s,x)+\gamma_{cv}(s,x)+\gamma_{cw}(s,x) \\ &\quad +\gamma_{hv}(s,x)+\gamma_{hw}(s,x)+\gamma_{dv}(s,x)+\gamma_{dw}(s,x)+\gamma_{r}(s,x), \end{aligned} \tag{12}$$

where

$$\gamma_{hw}(s,x)=\gamma_{hw0}(s,x)+\gamma_{hw1}(s,x)+\cdots+\gamma_{hwN}(s,x), \tag{13}$$

$$\gamma_{dw}(s,x)=\gamma_{dw0}(s,x)+\gamma_{dw1}(s,x)+\cdots+\gamma_{dwN}(s,x) \tag{14}$$

and

$$\gamma_{r}(s,x)=\gamma_{r}(s,x,1)+\gamma_{r}(s,x,2)+\cdots+\gamma_{r}(s,x,N). \tag{15}$$

In equations (12)-(15) flow *into* a state has been taken as positive and we require $s\neq x$. The total flow out of state s is denoted $q(s,s)$ and is given by

$$q(s,s)=-\sum_{\substack{k=0\\k\neq s}}^{S_{\max}} q(k,s). \tag{16}$$

The statistical equilibrium can be found using the flow balance equations that are set of $S_{\max}+1$ simultaneous equations for the unknown state probabilities,

$$\sum_{j=0}^{S_{\max}} q(i,j)\times p(j)=0, \qquad i=0,1,\ldots\ldots,S_{\max}-1 \tag{17}$$

$$\sum_{j=0}^{S_{\max}} p(j)=1, \tag{18}$$

in which, for, $i \neq j$, $q(i,j)$ is the net transition flow into state i from state j, and $q(i,i)$ is the total transition flow out of state i.

HAND-OFF ARRIVAL PARAMETERS

The average arrival rate of hand-off voice sessions, Λ_{hv}, the average hand-off arrival rate of active sessions, Λ_{hw}, the average hand-off arrival rate of k-suspended sessions, $\Lambda_r(k)$, the fraction of hand-off arrivals of voice sessions that are on g-type platform, F_{vg}, the fraction of hand-off arrivals of active data sessions that are from g-type platform, F_{wg}, and the fraction of hand-off arrivals of k-suspended sessions that are on g-type platform, $F_{rg}(k)$, can be determined from the dynamics of the process itself. An iterative method can be used [4]. The average hand-off departure rate of voice sessions on g-type platforms, $\Delta_{hv}(g)$, can be expressed as

$$\Delta_{hv}(g) = \sum_{s=0}^{S_{max}} \mu_D(g) \cdot v(s,g) \cdot p(s). \tag{19}$$

Thereafter, the overall average hand-off departure rate of voice sessions, Δ_{hv}, can be written as

$$\Delta_{hv} = \sum_{g=1}^{G} \Delta_{hv}(g). \tag{20}$$

The average hand-off departure rate of active data sessions on g-type platforms, $\Delta_{hw}(g)$, can be expressed as

$$\Delta_{hw}(g) = \sum_{s=0}^{S_{max}} \mu_D(g) \cdot w(s,g) \cdot p(s) \tag{21}$$

Thereafter, the overall hand-off departure rate of active data sessions, Δ_{hw}, can be written as

$$\Delta_{hw} = \sum_{g=1}^{G} \Delta_{hw}(g). \tag{22}$$

The average hand-off departure rates of k-suspended sessions on g-type platforms, $\Delta_r(g,k)$, can be expressed as

$$\Delta_r(g,k) = \sum_{s=0}^{S_{max}} \mu_D(g) \cdot r(s,g,k) \cdot p(s) \tag{23}$$

Also, the overall average hand-off departure rates of k-suspended sessions, $\Delta_r(k)$, can be written as

$$\Delta_r(k) = \sum_{g=1}^{G} \Delta_r(g,k). \tag{24}$$

From these equations, we find that the fraction of hand-off departures of voice sessions that are on g-type platforms, F'_{vg}, is

$$F'_{vg} = \Delta_{hv}(g) / \Delta_{hv}, \tag{25}$$

the fraction of hand-off departures of active data sessions that are on g-type platform, F'_{wg}, is

$$F'_{wg} = \Delta_{hw}(g) / \Delta_{hv}, \tag{26}$$

and, the fraction of hand-off departures of k-suspended sessions on g-type platforms, $F'_{rg}(k)$, is

$$F'_{rg}(k) = \Delta_r(g,k) / \Delta_r(k). \tag{27}$$

Since the maximum allowable reconnection attempts is N, there are N average hand-off departure rates and fractions, each corresponding to a value of k. Any hand-off departure of a g-type platform from a cell corresponds to a hand-off arrival to another cell. Therefore, the hand-off arrival and departure rates per cell for a homogeneous system in statistical equilibrium must be equal. That is we must have

$$\begin{aligned} F_{vg} &= F'_{vg} \\ F_{wg} &= F'_{wg} \\ F_{rg}(k) &= F'_{rg}(k) \\ \Lambda_{hv} &= \Delta_{hv} \\ \Lambda_{hw} &= \Delta_{hw} \\ \Lambda_r(k) &= \Delta_r(k) \end{aligned} \tag{28}$$

where $1 \le k \le N$.

PERFORMANCE MEASURES

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

Carried Traffic of Voice Sessions

An important performance measure from a system point of view is the carried traffic [4]. Since the traffic of data sessions is transparent to user of voice session, the carried traffic of voice sessions is same regardless of the amount of data sessions in the system. The carried traffic of voice sessions for g-type platforms, $A_{cv}(g)$, is

$$A_{cv}(g) = \sum_{s=0}^{S_{max}} v(s,g) \cdot p(s), \tag{29}$$

and the total carried traffic of voice sessions, A_{cv}, is

$$A_{cv} = \sum_{g=1}^{G} A_{cv}(g). \tag{30}$$

Carried Traffic of Data Sessions

Clearly, the traffic of data sessions strongly depends on the traffic of voice sessions in the system. The carried traffic of data sessions for g-type platform, $A_{cw}(g)$, is

$$A_{cw}(g) = \sum_{s=0}^{S_{max}} w(s,g) \cdot p(s), \tag{31}$$

and the total carried traffic of data sessions, A_{cw}, is

$$A_{cw} = \sum_{g=1}^{G} A_{cw}(g). \tag{32}$$

Average Number of k-Suspended Sessions

The average number of k-suspended sessions for g-type platforms, $A_{cr}(g,k)$, is

$$A_{cr}(g,k) = \sum_{s=0}^{S_{max}} r(s,g,k) \cdot p(s), \tag{33}$$

the average number of k-suspended sessions, $A_{cr}(k)$, is

$$A_{cr}(k) = \sum_{g=1}^{G} A_{cr}(g,k), \tag{34}$$

and, the average number of suspended sessions in a cell, A_{cr}, is

$$A_{cr} = \sum_{k=1}^{N} A_{cr}(k). \tag{35}$$

Blocking Probability of Voice Sessions

The blocking probability for voice sessions is the average fraction of newly generated voice sessions that are denied access to a channel. Since there is no quotas for specific type of mobility platform, the blocking probability is same for all type of platform. Blocking of newly generated voice sessions happens when the cell is on one of state of following disjoint subsets of states

$$L_{Bv} = \{s: J(s) \geq C - C_h\}. \tag{36}$$

And, the blocking probability of voice session, P_{Bv}, is expressed as

$$P_{Bv} = \sum_{s \in L_{Bv}} p(s). \tag{37}$$

Blocking Probability of Data Sessions

The blocking probability for data sessions is the average fraction of newly generated data sessions that are denied access to a channel. A newly generated data session will not be blocked unless all channels are occupied either by voice or data sessions. Blocking of newly generated data sessions happens when the cell is on one of state of following disjoint subsets of states

$$L_{Bw} = \{s: J(s) = C\}. \tag{38}$$

And, the blocking probability of data session, P_{Bw}, is expressed as

$$P_{Bw} = \sum_{s \in L_{Bw}} p(s). \tag{39}$$

Hand-Off Failure Probability of Voice Sessions

The hand-off failure probability of voice sessions, P_{Hv}, is the average fraction of hand-off attempts for voice sessions that are denied in the target cell because all channels are occupied by voice sessions in a cell. A failure of hand-off of a voice session happens when the cell is on one of state of following disjoint subsets of states

$$L_{Hv} = \{s: v(s) = C\}. \tag{40}$$

The hand-off failure probability of voice sessions is expressed as

$$P_{Hv} = \sum_{s \in L_H} p(s) . \tag{41}$$

Hand-Off Failure Probability of Data Sessions

The hand-off failure probability of data session, P_{Hw}, is the average fraction of hand-off attempts for data sessions that are denied in target cell because all channels are occupied *and N* suspended sessions are in the cell. A data session, either active or suspended, that are denied in the target cell due to the lack of system capacity will be forced into termination and cleared from the system. We define following disjoint set of states, in which hand-off attempts of data sessions will fail

$$L_{Hw} = \{s: v(s) = C, r(s) = H\}. \tag{42}$$

Then, the hand-off failure probability of data sessions is expressed as

$$P_{Hw} = \sum_{s \in L_{WH}} p(s) . \tag{43}$$

Forced Termination Probability of Voice Session

A voice session that fails its hand-off will be forced into termination. The forced termination probability of voice sessions on g-type platform, $P_{FTv}(g)$, is defined as the probability that a g-type voice session that is not blocked is interrupted due to hand-off failure during its life time as in [4], [8]. It can be shown that the forced termination probability of voice type session is given by

$$P_{FTv}(g) = \frac{\mu_D(g) \times P_{Hv}}{\mu_v(g) + \mu_D(g) \times P_{Hv}} . \tag{44}$$

Forced Termination Probability of Data Session

The forced termination probability of data session on g-type platform, $P_{FTw}(g)$, is defined as the probability that a data session that is not blocked is forced into termination during its lifetime. The prematurely terminated data session will be cleared from system. It should be reminded that there are three possible scenarios that a data session is forced into termination during its lifetime. Firstly, a data session, whether active or suspended, attempts its hand-off into a cell in which the system already supports maximum number of connections (either for voice sessions or data sessions) and maximum number of suspended sessions. In this case, a data session will forced to termination even if it hasn't finished its maximum allowable reconnection attempts. A premature termination of a data session, either active or suspended, that has not finished its maximum allowable reconnection attempts is called a ***non-maximum termination.*** Secondly, a suspended data session is forced terminated due to the failures of maximum number of reconnection attempts. It should be recalled that a suspended session attempts a hand-off when it moves to another cell even though the radio link between mobile unit and base was disconnected. The hand-off attempt generated by a suspended session will be counted as one reconnection attempt in the sense that the reconnection counter will be incremented. Therefore, the last attempt before termination can be either a retry attempt or a hand-off attempt. The termination of an N-suspended session due to the maximum number of allowable reconnection attempts being met is called a ***maximum termination***. Thirdly, If a voice session arrives when all channels are occupied *and* there are N suspended sessions and more than one active data session in the cell, an active data session must be terminated to accommodate the arriving voice session. The termination of an active data session due to an arrival of a voice session is called a ***pushed out termination***. Since the termination occurs while the mobile counter indicates less than N, this kind of termination is a ***non-maximum termination***.

A hand-off attempt of data session will succeed or fail according to the state of target cell. We define following disjoint and mutually exclusive set of states such that, in the set, a hand-off attempt of the data session will have same outcome.

$$L_A = \{s: v(s) < C\}. \tag{45}$$

$$L_D = \{s: v(s) = C, r(s) \neq H\}. \tag{46}$$

and L_{Hw}, that was defined in equation (42). When the target cell is in L_A, a hand-off attempt of the data session will succeed and, simply, a session will be active. When the cell is one of the state in L_D, the cell already supports maximum number of sessions (either of data or voice) *but* the number of suspended sessions in the cell is fewer than H. If a data session, either active or suspended, arrives due to a hand-off attempt when the target cell is one of these state, it will stay as a suspended session and wait for its reconnection opportunity unless the reconnection counter indicates maximum allowable reconnection attempts, $(N+1)$. When the reconnection counter indicates $(N+1)$, a session will be forced into termination. When the cell is one of the state in L_{Hw}, the all channels in target cell is occupied *and H* suspended sessions are in the cell. In this case, hand-off attempt of the data session will be failed and an arriving session will be prematurely terminated. The corresponding system probabilities is expressed as

$$P_A = \sum_{s \in L_A} p(s). \tag{47}$$

$$P_D = \sum_{s \in L_D} p(s). \tag{48}$$

Figure 1 is a flow graph depicting in the lifetime of a data session. As shown in the figure, there are two scenarios that an active data session to become a 1-suspended session. Firstly, an active data session try a hand-off attempt into a cell in which the state is in L_D (This is called "type 1 suspension"). The probability that an active data session on a g-type platform attempts a hand-off as a upcoming event before its session completes can be shown as $\mu_D(g)/[\mu_w(g)+\mu_D(g)]$ and the probability that the target cell is in L_D is simply P_D. Therefore, the probability, $\tau_{s1}(g)$, that an active data session on a g-type platform becomes a suspended session through a type 1 suspension can be expressed as

$$\tau_{s1}(g) = \frac{\mu_D(g) \cdot P_D}{\mu_w(g) + \mu_D(g)}. \tag{49}$$

Secondly, an active data session is suspended due to an arrival of a voice session when a cell state is in L_D (This is called "type 2 suspension"). There are nine possible state dependent driving processes when a cell state is s. Those are listed in section "DRIVING PROCESSES". In these driving processes, only a generation of voice session and a hand-off arrivals of voice session can cause a type 2 suspension of active data session. We define following disjoint set of states for explanation of these driving processes.

$$L_1 = \{s: J(s) = C, 0 \le v(s) < C_h, r(s) \ne H\}, \tag{50}$$

$$L_2 = \{s: J(s) = C, C_h \le v(s) < C, r(s) \ne H\}. \tag{51}$$

When the cell is in L_1, all processes can potentially occur as a upcoming event. However, a generation of voice session and hand-off arrival of voice session can cause a type 2 suspension. The probability, P_1, that either a generation of voice session or a hand-off arrival of voice session will be the next upcoming event, when the cell is in state s, can be expressed as

$$P_1 = \sum_{g=1}^{G} \frac{\Lambda_{nv}(g) + \Lambda_{hv} \cdot F_{vg}}{D_1}, \qquad s \in L_1 \tag{52}$$

where

$$\begin{aligned} D_1 = {} & \Lambda_{nv}(g) + \Lambda_{nw}(g) + (\mu_v(g) \times v(s,g)) + (\mu_w(g) \times w(s,g)) \\ & + (\Lambda_{hv} \times F_{vg}) + (\Lambda_{hw} \times F_{wg}) + (\Lambda_r(k) \times F_{rg}) + (\mu_D(g) \times v(s,g)) \\ & + (\mu_D(g) \times w(s,g)) + (\mu_D(g) \times r(s,g)) + \sum_{k=1}^{N} (\mu_r(k,g) \times r(x,g,k)). \end{aligned} \tag{53}$$

When the cell is one of the state in L_2, all possible processes *except* generation of voice sessions can occur as a next upcoming event. However, *only* a hand-off arrival of voice session cause a type 2 suspension. Therefore, the probability, P_2, that a hand-off arrival of voice session will be the next upcoming event, when the cell is in state s, can be expressed as

$$P_2 = \sum_{g=1}^{G} \frac{\Lambda_{hv} \cdot F_{vg}}{D_2}, \qquad s \in L_2 \tag{54}$$

where

$$\begin{aligned} D_2 = {} & \Lambda_{nw}(g) + (\mu_v(g) \times v(s,g)) + (\mu_w(g) \times w(s,g)) \\ & + (\Lambda_{hv} \times F_{vg}) + (\Lambda_{hw} \times F_{wg}) + (\Lambda_r(k) \times F_{rg}) + (\mu_D(g) \times v(s,g)) \\ & + (\mu_D(g) \times w(s,g)) + (\mu_D(g) \times r(s,g)) + \sum_{k=1}^{N} (\mu_r(k,g) \times r(x,g,k)). \end{aligned} \tag{55}$$

Thereafter, the conditional probability, $\tau_{s2}(g)$, that an active data session on a g-type platform becomes a suspended session through a type 2 suspension can be expressed as

$$\tau_{s2}(g) = \begin{cases} Q_s(s,g) \times P_1, & \text{if} \quad s \in L_1 \\ Q_s(s,g) \times P_2, & \text{if} \quad s \in L_2 \\ 0 & \textit{otherwise.} \end{cases} \tag{56}$$

Consequently, the probability, $P_\xi(g)$, that an active data session on a g-type platform becomes a 1-suspended session due to either type 1 suspension or type 2 suspension can be expressed as

$$P_\xi(g) = \tau_{s1}(g) + \tau_{s2}(g). \quad (57)$$

A suspended data session will attempt to reconnect to base station. This reconnection process can succeed or fail according to the cell state. For a suspended session, there are two possible upcoming events. When the dwell time of suspended session is less than the next trial time, the session will attempt hand-off. Otherwise, the session will attempt retry. The probability that a g-type k-suspended session attempts hand-off as a next upcoming event can be written as $\mu_D(g)/[\mu_r(k,g) + \mu_D(g)]$. Also, the probability that a g-type k-suspended session attempts a retry as the next upcoming event can be written as $\mu_r(k,g)/[\mu_r(k,g) + \mu_D(g)]$.

A hand-off event of a suspended session, can result in resumption of an active session, a forced termination, or continued suspension (with an incremented reconnection attempt counter). These events occur with respective P_A, P_{Hw}, and P_D.

A *retry* attempt generated by a suspended session can result in: 1) the suspended session becoming active. (This event occurs with probability, P_A. See equation (47).) ; or, 2) the suspension of the session being continued (with an incremented reconnection counter if the maximum number of allowed reconnection attempts has not been reached (This event occurs with probability $P_{Hw} + P_D$. See equation (43) and (48).) ; or, 3) termination of the suspended session. (This event occurs if the retry attempt fails and the maximum number of reconnection attempts is exhausted).

A g-type suspended data session will become active if a *reconnection* attempt (retry or hand-off) is successful. The probability of this event is simply P_A

There are two possible scenarios in which a k-suspended session can become a $(k+1)$-suspended session. One is to attempt a hand-off and be continued suspension. The other is to try a retry attempt and be continued suspension. Therefore, the probability, $\theta_k(g)$, that a k-suspended session on a g-type platform becomes a $(k+1)$-suspended session through either a retry of a hand-off attempt can be written as

$$\theta_k(g) = \frac{\mu_r(k,g)\cdot(P_{Hw}+P_D)}{\mu_r(k,g)+\mu_D(g)} + \frac{\mu_D(g)\cdot P_D}{\mu_r(k,g)+\mu_D(g)} \tag{58}$$

$$= \frac{\mu_r(k,g)\cdot(P_{Hw}+P_D)+\mu_D(g)\cdot P_D}{\mu_r(k,g)+\mu_D(g)} \tag{59}$$

where $1 \leq k < N$.

Consider an *active* data session. An active session may be suspended. This occurs if it undergoes an unsuccessful hand-off when there is still room for suspended calls, or if the active session is preempted by a voice call while there is still waiting room for suspended calls. A *suspended* session undergoes reconnection attempts each of which can result in: 1), the *suspended* session being continued in suspension; 2) the *suspended* session becoming active; or, 3) the *suspended* session being terminated. A *suspended* session will become active again if it undergoes either *a successful hand-off or a successful reconnection attempt.* Let $\eta(g)$, be the probability, that an *active* data session on a g-type platform either remains active or becomes active again (after undergoing either a hand-off attempt or a preemption by an arriving voice session). Then, if we define $\theta_0(g)=1$, we find

$$\eta(g) = \left(\frac{\mu_D(g)\cdot P_A}{\mu_w(g)+\mu_D(g)}\right) + (P_\xi(g)\cdot P_A) + (P_\xi(g)\cdot P_A\cdot\theta_1(g))$$

$$+ (P_\xi(g)\cdot P_A\cdot\theta_1(g)\cdot\theta_2(g)) + \ldots\ldots\ldots + (P_\xi(g)\cdot P_A\cdot\theta_1(g)\cdots\theta_{N-1}(g)). \tag{60}$$

$$= \frac{\mu_D(g)\cdot P_A}{\mu_w(g)+\mu_D(g)} + P_\xi(g)\cdot P_A\cdot\sum_{k=0}^{N-1}\prod_{m=0}^{k}\theta_m(g) \tag{61}$$

The first term of the RHS on (61) is the probability that an active session undergoes a successful hand-off attempt. The second term is the probability that an active session is suspended, remains suspended through (*k-1*) reconnection attempts, and succeeds (becomes active) on its k^{th} attempt.

NON-MAXIMUM TERMINATION PROBABILITY

In previous discussion, we defined the premature termination of a data session, either active or suspended, as one which has not exhausted its maximum allowable number of reconnection attempts as a ***non-maximum termination***. Therefore, if the reconnection counter indicates ***less than*** (*N+1*) when a session is forced to terminate, the termination is a non-maximum termination. We define the following disjoint set of states to explain this process in more detail. Let

$$L_3 = \{s\colon J(s) = C, 0 \le v(s) < C_h, r(s) = H\}, \tag{62}$$

$$L_4 = \{s\colon J(s) = C, C_h \le v(s) < C, r(s) = H\}. \tag{63}$$

When the cell is in a state belonging to L_3, any processes can potentially generate the next event. However, the arrival of a new voice session *or* the hand-off arrival of voice session can cause a pushed out termination. In this region, the probability, P_3, that either a generation of voice session or a hand-off arrival of voice session will be the next upcoming event, when the cell is in state *s*, can be expressed as

$$P_3 = \sum_{g=1}^{G} \frac{\Lambda_{nv}(g) + \Lambda_{hv} \cdot F_{vg}}{D_1}, \qquad s \in L_3 \tag{64}$$

where D_1 is given by equation (53).

When the cell is in a state belonging to L_4, any process *except* generation of new voice sessions can cause the next state transition. However, *only* the hand-off arrival of a voice session can cause a pushed out termination. The probability, P_4, that a hand-off arrival of a voice session will be the next upcoming event, when the cell is in state *s*, can be expressed as

$$P_4 = \sum_{g=1}^{G} \frac{\Lambda_{hv} \cdot F_{vg}}{D_2}, \qquad s \in L_4 \tag{65}$$

where D_2 is expressed in equation (55).

Thereafter, the conditional probability, $\xi(g)$, that an active data session on a g-type platform is terminated due to pushed out termination can be expressed as

$$\xi(g) = \begin{cases} Q_t(s,g) \times P_3, & if \quad s \in L_3 \\ Q_t(s,g) \times P_4, & if \quad s \in L_4 \\ 0 & otherwise. \end{cases} \tag{66}$$

The non-maximum termination probability, $P_{NT}(g)$, is defined as the probability that a data session on g-type platform, either active or suspended, is forced into termination during its lifetime even if it has not finished its maximum allowable reconnection attempts. It should be noted that a data session can be terminated (non-maximum termination) as an active session, a 1-suspended session, a 2-suspended session, ……, or *N-1* suspended session.

The probability, $\varphi_\xi(g)$, that an active g-type data session is pushed out into termination due to an arrival of voice session during its lifetime can be written as

$$\varphi_\xi(g) = \xi(g) + \eta(g) \cdot \xi(g) + \eta^2(g) \cdot \xi(g) + \cdots\cdots \tag{67}$$

$$= \xi(g) \cdot \sum_{i=0}^{\infty} (\eta(g))^i \tag{68}$$

The probability, $\varphi_0(g)$, that a g-type suspended data session is forced into termination (while the session is in progress) *due to a failure of hand-off* can be written as

$$\varphi_0(g) = \frac{\mu_D(g) \cdot P_{Hw}}{\mu_D(g) + \mu_w(g)} + \frac{\mu_D(g) \cdot P_{Hw} \cdot \eta(g)}{\mu_D(g) + \mu_w(g)} + \frac{\mu_D(g) \cdot P_{Hw} \cdot \eta^2(g)}{\mu_D(g) + \mu_w(g)} + \cdots\cdots \tag{69}$$

$$= \frac{\mu_D(g) \cdot P_{Hw} \cdot \sum_{i=0}^{\infty} (\eta(g))^i}{\mu_D(g) + \mu_w(g)}. \tag{70}$$

A failure of hand-off attempt will cause a termination of 1-suspended session if the state of target cell is L_3 or L_4. It should be noted that a failure of retry attempt will not cause termination of k-suspended data session, unless $k = N$. The probability, $\varphi_1(g)$, that a g-type suspended data session is forced into termination while it is a 1-suspended session due to a failure of hand-off can be written as

$$\varphi_1(g) = P_\xi(g) \cdot \sum_{i=0}^{\infty} (\eta(g))^i \cdot \left(\frac{\mu_D(g) \cdot P_{Hw}}{\mu_D(g) + \mu_r(1, g)} \right). \tag{71}$$

With the same fashion, the probability, $\varphi_k(g)$, that a data session on a g-type platform is forced into termination while it is a k-suspended session, where $0<k<N$, *due to a failure of hand-off* can be written as

$$\varphi_k(g) = P_\xi(g) \cdot \sum_{i=0}^{\infty} (\eta(g))^i \cdot \left(\frac{\mu_D(g) \cdot P_{Hw}}{\mu_D(g) + \mu_r(1, g)} \right) \cdot \prod_{m=1}^{k-1} \theta_m(g). \tag{72}$$

Then, the overall non-maximum probability, $P_{NT}(g)$, can be written as

$$P_{NT}(g) = \varphi_\xi(g) + \varphi_0(g) + \varphi_1(g) + \cdots + \varphi_{N-1}(g) \tag{73}$$

$$= \sum_{i=0}^{\infty} (\eta(g))^i \cdot \left[\xi(g) + P_{Hw} \cdot \left[\frac{\mu_D(g)}{\mu_D(g) + \mu_w(g)} + \frac{\mu_D(g) \cdot P_\xi(g)}{\mu_D(g) + \mu_r(k, g)} \cdot \sum_{k=1}^{N-1} \prod_{m=0}^{k-1} \theta_m(g) \right] \right] \tag{74}$$

$$= \frac{1}{1 - \eta(g)} \cdot \left[\xi(g) + P_{Hw} \cdot \left[\frac{\mu_D(g)}{\mu_D(g) + \mu_w(g)} + \frac{\mu_D(g) \cdot P_\xi(g)}{\mu_D(g) + \mu_r(k, g)} \cdot \sum_{k=1}^{N-1} \prod_{m=0}^{k-1} \theta_m(g) \right] \right]. \tag{75}$$

MAXIMUM TERMINATION PROBABILTITY

The maximum termination probability, $P_{MT}(g)$, is defined as the probability that a data session on a g-type platform is forced into termination during its lifetime because the maximum allowable number of reconnection attempts has been reached. The last attempt before termination can be either a hand-off or retry. The maximum termination probability, $P_{MT}(g)$, can be written as

$$P_{MT}(g)=\sum_{i=0}^{\infty}\eta(g)\cdot P_{\xi}(g)\cdot(P_{Hw}+P_D)\cdot\prod_{i=1}^{N-1}\theta_i(g) \quad (76)$$

$$=\frac{1}{1-\eta(g)}\cdot P_{\xi}(g)\cdot(P_{Hw}+P_D)\cdot\prod_{i=1}^{N-1}\theta_i(g). \quad (77)$$

Thereafter, the total forced termination probability of data session, $P_{FTw}(g)$, can be written as

$$P_{FTw}(g)=P_{NT}(g)+P_{MT}(g). \quad (78)$$

Average Time Per Suspension

The average time per suspension, $W(g)$, is the expected amount of time that reconnection process will carry on after a data session to be suspended. Therefore, it is the average time frame from the point that an active data session is suspended to the point that a suspended session is active or is forced into termination.

To calculate the average time per suspension, determine the average rate of suspension and the average number of suspended sessions using state probabilities. Little's law will then be applied to find the average time per suspension [9].

As we discussed before, there are two types of suspension: type 1 suspension and type 2 suspension. The average rate of suspension of sessions on g-type platform, $H(g)$, is given by

$H(g)=$ the rate of type 1 suspension + the rate of type 2 suspension.

$$H(g)=\left(\sum_{s\in L_D}p(s)\cdot\Lambda_{hv}\cdot F_{vg}\right)+\left[\left(\sum_{s\in L_1}p(s)\cdot\sum_{g=1}^{G}[\Lambda_{nv}(g)+\Lambda_{hv}\cdot F_{vg}]\right)+\left(\sum_{s\in L_2}p(s)\cdot\sum_{g=1}^{G}\Lambda_{hv}\cdot F_{vg}\right)\right] \quad (79)$$

$$=\left(P_D\cdot\Lambda_{hv}\cdot F_{vg}\right)+\left(P_1\cdot\sum_{g=1}^{G}[\Lambda_{nv}(g)+\Lambda_{hv}\cdot F_{vg}]\right)+\left(P_2\cdot\Lambda_{hv}\right). \quad (80)$$

We can determine the average number of suspended sessions from g-type platforms, $A_{cr}(g)$, from equation (33) as follows

$$A_{cr}(g) = \sum_{k=1}^{N} A_{cr}(g,k). \tag{81}$$

Thereafter, we can find the average time per suspension from g-type data session, $W(g)$. Using Little's law, this is

$$W(g) = A_{cr}(g) / H(g). \tag{82}$$

Average Number of Suspensions Per Data Session

When data session arrives, it may be accommodated or blocked according to the state of a cell. So, we can expressed the average accommodation rate of data sessions on g-type platform in a cell, $E(g)$, as follows

$$E(g) = \Lambda_{nw}(g) \cdot (1 - P_{Bw}). \tag{83}$$

Thereafter, the average number of suspensions per session, $S(g)$, can be written as

$$S(g) = H(g) / E(g). \tag{84}$$

Average Time in Suspension Per Data Session

The average time in suspension per data session on g-type platform, $M_s(g)$, is the expected amount of time that a data session is suspended during its lifetime. Using Little's law, this can be written as

$$M_s(g) = S(g) \cdot W(g) = \frac{H(g)}{E(g)} \cdot \frac{A_{cr}(g)}{H(g)} = \frac{A_{cr}(g)}{E(g)}. \tag{85}$$

Average Total Lifetime of a Data Session

The average total lifetime of a data session, regardless how it ends (successfully complete or be prematurely terminated) on a g-type platform, $M_t(g)$, is the expected time that a data session spends in the system. Using Little's law, this can be written as

$$M_t(g) = \frac{A_{cr}(g) + A_{cw}(g)}{E(g)}. \tag{86}$$

Suspension Time of Data Session

The fraction of a data session's lifetime remaining to be suspended is an important performance metric. For a data session on a g-type platform, this is denoted as $L(g)$. It can be written as

$$L(g)=\frac{M_s(g)}{M_t(g)}=\frac{A_{cr}(g)}{A_{cw}(g)+A_{cr}(g)}. \tag{87}$$

DISCUSSION OF RESULTS

Numerical results were generated using the approach described in this paper. For Figure 3-11, an unencumbered voice session duration of 100s was assumed and an unencumbered data session duration of 20s was assumed. Two platform types, low mobility and high mobility, were considered. The mean dwell time of 500s was assumed for a low mobility platform and that of 100s was assumed for a high mobility platform. A homogeneous system was assumed. The mean k-trial time of a g-type k-suspended session was chosen to be 10s for $1 \le k \le N$ (That is, $\overline{T}_r(k,g)=10s$).

The number of possible states depends on the number of platform types (G), the number of channels (C), the number of allowable suspended sessions in a cell (H), and the number of allowable reconnection attempts (N). For example, the number of possible states with parameters used to generate Figure 6 (G=2, C=15, H=3, N=3) is 325,584. All calculations were made using Sun Ultra-2 and Sun Ultra-1 workstations. For the parameters of Figure 6, the cpu time needed to obtain the solution is about 92 min. using Sun Ultra-2 workstation.

The abscissas for Figures 3-11 reflect call demands with the assumptions stated above. In each of these, the abscissa is the new voice session origination rate for platform type 1 (denoted $\Lambda_{nv}(1)$). The ratio of new voice session generation rates from other platform types to that of type 1 platforms were held fixed with parameters $\alpha(g)$. Also, the new data session generation rate for platform type g is determined with respect to new voice session origination rate using parameters, $\beta(g)$. For all calculations in this paper, we used $\alpha(g)=\beta(g)=1$.

Figure 3-5 shows voice traffic performances. Since the traffic of data sessions is transparent to users of voice session, the traffic performance of voice sessions with data traffic is identical to that without data traffic. Figure 3 shows the blocking probability of voice session. As the number of reserved channels, C_h, is increased, obviously more newly generated voice session can not be accommodated. For an abscissa value of 1.3×10^{-4} calls/sec, blocking probability increases by one and half orders of magnitude as C_h changes from 0 to 4.

Figure 4 shows the forced termination probability of voice session. As the number C_h increase, fewer voice sessions are forced into termination with the cost of blocking of more newly generated voice sessions. Over the whole traffic range, forced termination probability of a voice session decreases by an order of magnitude each time C_h is increased by 2 channels. Voice sessions on fast mobile platforms have higher forced termination probability than that on slow mobile platform. This is because voice sessions on fast mobile platforms are more likely to experience more hand-offs during lifetime of a session. It is also seen that the decease in forced termination probability (as C_h increases) is much greater than the corresponding increase in blocking probability. This allows a very favorable exchange.

Figure 5 shows the carried traffics of voice sessions. As expected, it is seen that if more channels are reserved for hand-off attempts, there is less carried traffic. This is because more channels are reserved for hand-off attempts, the more newly generated voice sessions will be blocked. In the traffic range shown, it is seen that increasing C_h from 2 to 4 has a more pronounced effect on performance than increasing C_h from 0 to 2. For example, for an abscissa value of 1.6×10^{-4} calls/sec, increasing C_h from 0 to 2 decreases carried traffic of voice sessions from 8.6 erlangs to 8.4 erlangs (about 2.3 %), but that from 2 to 4 decreases carried traffic of voice sessions from 8.4 erlangs to 7.9 erlangs (about 6%).

Figure 6 shows the average number of data sessions per cell, either active or suspended. Figure 7 shows the blocking probability of data sessions. It is seen that if more channels are reserved for hand-off attempts of voice sessions, more data sessions can be accommodated. This is because the more channels are reserved for hand-off attempts, the more newly generated voice sessions will be blocked. Consequently, the opportunity for data session to be accommodated is increased. In this traffic range, it is seen that increasing C_h from 2 to 4 cause more change of performance than increasing C_h from 0 to 2. For an abscissa value of 1.6×10^{-4} calls/sec, increasing C_h from 0 to 2 increases carried traffic of data sessions from 1.83 erlangs to 1.84 erlangs (about 0.5 %), but that from 2 to 4 increases carried traffic of data sessions from 1.84 erlangs to 1.88 erlangs (about 2.1%).

Figure 8 shows the forced termination probability of data sessions for various values of C_h used for cut-off priority. When we increase C_h, clearly, forced termination probability of data sessions decreased. Over the whole traffic range, forced termination probability of data session decreases by an order of magnitude each time C_h is increased by 2 channels. It is also seen that data sessions on slow mobiles have smaller forced termination probability. This is because a data session on a slow mobile is more likely to finish its session with fewer hand-offs during its lifetime.

Figure 9 shows the dependence of forced termination probability of data sessions on the number of maximum allowable reconnection attempts, *N*. As we can see, increasing *N* results in fewer data sessions being forced to terminate during their lifetime. For an abscissa value of 1.6×10^{-4} calls/sec, forced termination probability of data session decreases by two orders of magnitude as *N* increases from 1 to 3.

Figure 10 shows the dependence of forced termination probability of data sessions on the number of maximum supportable suspended sessions, *H*. With increasing *H*, clearly, the less the forced termination probability of data session is expected. For the same abscissa value with last paragraph, (1.6×10^{-4} calls/sec), forced termination probability of data session decreases by an order of magnitude as *H* increases from 2 to 3.

Figure 11 shows the dependence of average number of suspensions per data session, $S(g)$, on traffic demands. Consider suspended data sessions on fast mobiles in comparison with those on slow platforms. It is seen that data sessions on slow mobile have smaller number of suspension during a session. For example, for an abscissa value of 1.6×10^{-4} calls/sec, the expected number of suspensions for a data session on a slow mobile is 0.08, however, for fast mobile it is 0.09. This is because a data session on slow mobile is more likely to finish its session with fewer hand-offs during its lifetime, and, therefore, fewer suspensions are expected. It is also seen that with increasing N, the average number of suspensions per data session increases. This is *not* because data session will experience more suspensions given time *but* because the lifetime of a session is increased with increasing N.

CONCLUSIONS

With rapidly growing interest in the area of multimedia and mobile computing, the issue "how to accommodate diverse traffic types in wireless network" must be solved. In this paper, we propose a scheme in which each type of media is managed with different strategy according to the characteristics. For time insensitive data sessions, the system allows users to continue in a temporary off-line mode while awaiting an active network connection in the background. For time sensitive voice session, the system give preemptive priority over data traffic so that the transparency of data traffic is guaranteed to voice users. We develop the framework, using a suitable state description and multidimensional birth-death process representation, to compute traffic performance characteristics of proposed system. Consideration of mixed platform mobilities is accommodated in the model. Traffic performance depends strongly on the amount of priority given for hand-off attempt of voice sessions and the limit on the number of allowable reconnection attempts for data sessions. Relevant traffic performance measures can be calculated for given system parameters.

REFERENCES

[1] CCITT Recommendation I.212, "Broadband Aspects of ISDN," CCITT IXth Plenary Assembly, Melbourne, Australia, vol. 3, fasc. III. 7, 1998

[2] T. Saadawi, M. Ammar and A. Hakeem, Fundamentals of Telecommunication Networks, New York: John Wiley & Sons, Inc. 1994.

[3] D. Hong and S. S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile

Radio Telephone Systems with Prioritized and Non-Prioritized Hand-off Procedures," IEEE Trans. on Vehicular Technology, Aug. 1986, vol.VT-35, no. 3, pp. 77-92.

[4] S. S. Rappaport, "Blocking Hand-off and Traffic Performance Analysis for Cellular Communication System with Mixed Platforms," IEE (British) Proceedings, Part I, Communications, Speech and Vision, October 1993, vol. 40, no. 5, pp. 389-401.

[5] C. Purzynski and S. S. Rappaport, "Prioritized Resource Assignment for Mobile Cellular Communication Systems with Mixed Services and Platform Types," IEEE Trans. on Vehicular Technology, Aug. 1996, vol. 45, no. 3, pp. 443-458.

[6] W. D. Grover, W. A. Krzymien, J. C. Chin, "Background Data Service for Cellular Mobile Radio," in *Proc. 43rd IEEE Vehic. Technol. Conf.*, pp. 766-769, 1993.

[7] W. D. Grover, W. A. Krzymien, J. C. Chin, "Strategies for Management of Connection-Oriented Data Sessions Employing the Idle Times in Cellular Telephony," IEEE Trans. on Vehicular Technology, May 1995, vol. 44, no. 2, pp. 244-252.

[8] Y. Park and S. S. Rappaport, "Cellular Communication Systems with Voice and Background Data," in D. J. Goodman and D. Raychaudhuri (Eds.) 'Mobile Multimedia Communication,' (Plenum Press, New York, 1997), pp. 33-42.

[9] Y. Park and S. S. Rappaport, "Performance Analysis of Session Oriented Data Communications for Mobile Computing in Cellular Systems," CEAS Technical Report no. 744, Aug. 4, 1997, College of Engineering and Applied Science, State University of New York, Stony Brook, NY 11794

[10] S. S. Rappaport, "The Multiple-Call Hand-off Problem in High-Capacity Cellular Communication Systems," IEEE Trans. on Vehicular Technology, Aug. 1991, vol. VT-40, no. 3, pp. 546-557.

[11] P. V. Orlik and S. S. Rappaport, "Traffic Performance and Mobility Modeling of Cellular Communications with Mixed Platforms and Highly Variable Mobilities," *Proc. IEEE Vehic. Technol. Conf.*, pp. 587-591, 1997.

[12] P. V. Orlik and S. S. Rappaport, "A Model for Teletraffic Performance and Channel Holding Time Characterization in Wireless Cellular Communications with General Session and Dwell Time Distributions," CEAS Technical Report no. 742, July 1997, College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794.

[13] Y. Park and S. S. Rappaport, "Session Oriented Data Communications for Mobile Computing in Cellular Systems with Multiple Traffic Classes," CEAS Technical Report no. 760, Appendix B. College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794.

APPENDIX *A*: *Operation of the Reconnection Counter at Mobile Terminal.*

Figure 2 shows the detailed operation of the reconnection counter and reconnection process. When a mobile platform with an active data session moves to a target cell that has no channels available, the session will be suspended and the mobile's reconnection counter will be set to 1 if there are fewer than H suspended sessions already in the target cell (type 1 suspension). In addition, an active data type session will be suspended due to an arrival of preemptive voice session when a cell state is in L_D (type 2 suspension). In the target cell, additional reconnection attempts may be made. The mobile's counter will be incremented for each unsuccessful attempt. It may happen that the supporting platform moves to yet another cell. At that time a hand-off attempt will be made. This hand-off attempt will count as an attempted reconnection. Specially, if the hand-off attempt succeeds in getting a channel, the session will be counted and reconnection counter will be set to 0. If there are no channels in the target cell but there are fewer than H suspended sessions the reconnection effort will continue (this is, the counter will be incremented to 2 in this example case). If there are no channels available and there are H suspended sessions in the new target cell, the data session will be terminated. For non-terminated data sessions, the process will continue in this way as long as there has not been N consecutive failed reconnection (hand-off or retry) attempts. When this limit is reached the session will be forced to terminate. The mobile terminal may confirm this termination to the network via the control channel (A timeout in the network can also be used as a backstop).

APPRNDIX *B*: *Driving Processes and State Transition Rates.*

For convenience, dummy variable i and j are used as follow: $i = 1, 2, 3, \dots \dots, G$; $j = 1, 2, 3, \dots\dots\dots, N$.

Generation of Voice Sessions

Because of cut-off priority, C_h connections are held for arrivals of hand-off attempts (either of voice session or date session) or reconnection attempts (of suspended data session). Therefore, when the number of active sessions, either of voice or data, is fewer then C *and* the number of voice sessions is less than C_h in the cell, an arriving voice session will cause the state variable $v(x_{nv}, g)$ to be incremented by 1. Thus a permissible state x_{nv} is a predecessor state of s for a new voice session arrives on a g-type platform, if $J(x_{nv}) < C$, $v(x_{nv}) < C_h$, and the state variables are related by

$$\begin{aligned} &v(x_{nv,}g) = v(s, g) - 1 \\ &v(x_{nv}, i) = v(s, i), \qquad i \neq g \\ &w(x_{nv}, i) = w(s, i) \\ &r(x_{nv}, i, j) = r(s, i, j). \end{aligned} \tag{88}$$

When a cell has C active sessions (either of voice or data), fewer than C_h connections are occupied by voice sessions, and fewer than H suspended data sessions exist in the cell, an arriving voice session make an active data session be suspended. Suppose a data session on a m-type platform is suspended due to an arrival of a voice session on a g-type platform, then state variables are related by

$$
\begin{aligned}
&v(x_{nv}, g) = v(s, g) - 1 \\
&v(x_{nv}, i) = v(s, i), \quad i \neq g \\
&w(x_{nv}, m) = w(s, m) + 1 \\
&w(x_{nv}, i) = w(s, i), \quad i \neq m \\
&r(x_{nv}, m, 1) = r(s, m, 1) - 1 \\
&r(x_{nv}, m, j) = r(s, m, j), \quad j \neq 1 \\
&r(x_{nv}, i, j) = r(s, i, j), \quad i \neq m,
\end{aligned}
\tag{89}
$$

if $J(x_{nv}) = C$, $v(x_{nv}) < C_h$, and $r(x_{nv}) < H$.

If a voice session arrives when a cell has C active sessions (either of voice or data), fewer than C_h connections are for voice sessions, and fewer than H suspended data sessions are in the cell, an arriving voice session make an active data session be terminated. Suppose an active data session on a m-type platform is terminated due to an arrival of a voice session on a g-type platform, then state variables are related by

$$
\begin{aligned}
&v(x_{nv}, g) = v(s, g) - 1 \\
&v(x_{nv}, i) = v(s, i), \quad i \neq g \\
&w(x_{nv}, m) = w(s, m) + 1 \\
&w(x_{nv}, i) = w(s, i), \quad i \neq m \\
&r(x_{nv}, i, j) = r(s, i, j),
\end{aligned}
\tag{90}
$$

if $J(x_{nv}) = C$, $v(x_{nv}) < C_h$, and $r(x_{nv}) = H$.

Let $\Lambda_{nv}(g)$ denote the average arrival rate per cell of new voice session from g-type platforms. Then, the corresponding transition flow is given by

$$
\gamma_{nv}(s, x_{nv}, g) = \begin{cases} \Lambda_{nv}(g), & J(x_{nv}) < C \\ Q_s(x_{nv}, g) \times \Lambda_{nv}(g), & J(x_{nv}) = C, v(x_{nv}) < C_h, r(x_{nv}) < H \\ Q_t(x_{nv}, g) \times \Lambda_{nv}(g), & J(x_{nv}) = C, v(x_{nv}) < C_h, r(x_{nv}) = H. \end{cases}
\tag{91}
$$

Generation of Data Sessions

When there are fewer than C active sessions in a cell, an arriving data session can obtain a connection. Thus a permissible state x_{nw} is a predecessor state of s for an arrival of data session on a g-type platform, if $J(x_{nw}) < C$, and the state variables are related by

$$\begin{aligned} &v(x_{nw}, i) = v(s, i) \\ &w(x_{nw}, g) = w(s, g) - 1 \\ &w(x_{nw}, i) = w(s, i), \qquad i \neq g \\ &r(x_{nw}, i, j) = r(s, i, j) \end{aligned} \tag{92}$$

Let $\Lambda_{nw}(g)$ denote the average arrival rate per a cell of new data session from g-type platforms. Then, the corresponding transition flow is given by

$$\gamma_{nw}(s, x_{nw}, g) = \Lambda_{nw}(g). \tag{93}$$

Completion of Voice Sessions

A transition into state s, due to successful completion of a voice session on a g-type platform when the cell is in state x_{cv}, will cause the state variable $v(x_{cv}, g)$ to be decreased by 1. Thus a permissible state x_{cv} is a predecessor state of s for completion of a voice session on a g-type platform if the state variables are related by

$$\begin{aligned} &v(x_{cv}, g) = v(s, g) + 1 \\ &v(x_{cv}, i) = v(s, i), \qquad i \neq g \\ &w(x_{cv}, i) = w(s, i) \\ &r(x_{cv}, i, j) = r(s, i, j). \end{aligned} \tag{94}$$

The transition flow into state s from x_{cv} due to a completion of a voice session is given by

$$\gamma_{cv}(s, x_{cv}, g) = \mu_v(g) \times v(x_{cv}, g). \tag{95}$$

Completion of Data Sessions

A transition into state s, due to successful completion of a data session on a g-type platform when the cell is in state x_{cw}, will cause the state variable $w(x_{cw}, g)$ to be decreased by 1. Thus a permissible state x_{cw} is a predecessor state of s for completion of a data session on a g-type platform if the state variables are related by

$$
\begin{aligned}
&v(x_{cw}, i) = v(s, i) \\
&w(x_{cw}, g) = w(s, g) + 1 \\
&w(x_{cw}, i) = w(s, i), \qquad i \neq g \\
&r(x_{cw}, i, j) = r(s, i, j).
\end{aligned} \tag{96}
$$

The transition flow into state s from x_{cw} due to a completion of a data session is given by

$$\gamma_{cw}(s, x_{cw}, g) = \mu_w(g) \times w(x_{cw}, g). \tag{97}$$

Hand-Off Arrival of Voice Sessions

When a cell has fewer than C voice sessions *and* there are, at least, one available empty channel in a cell, an arriving voice session due to a hand-off can obtain a connection without making an active data session be suspended or be terminated. Thus a permissible state x_{hv} is a predecessor state of s for a hand-off arrival of a voice session on a g-type platform, if $J(x_{hv}) < C$, and the state variables are related by

$$
\begin{aligned}
&v(x_{hv}, g) = v(s, g) - 1 \\
&v(x_{hv}, i) = v(s, i), \qquad i \neq g \\
&w(x_{hv}, i) = w(s, i) \\
&r(x_{hv}, i, j) = r(s, i, j).
\end{aligned} \tag{98}
$$

When a cell support C connections (either for voice or data sessions) *and* fewer than H suspended data sessions exist in a cell, an arriving voice session due to a hand-off attempt make an active data session be suspended. Suppose a data session on a m-type platform is suspended due to a hand-off arrival of a voice session on a g-type platform, then state variables are related by

$$
\begin{aligned}
&v(x_{hv}, g) = v(s, g) - 1 \\
&v(x_{hv}, i) = v(s, i), \qquad i \neq g \\
&w(x_{hv}, m) = w(s, m) + 1 \\
&w(x_{hv}, i) = w(s, i), \qquad i \neq m \\
&r(x_{hv}, m, 1) = r(s, m, 1) - 1 \\
&r(x_{hv}, m, j) = r(s, m, j), \qquad j \neq 1 \\
&r(x_{hv}, i, j) = r(s, i, j), \qquad i \neq m
\end{aligned} \tag{99}
$$

if $J(x_{hv}) = C$, $r(x_{hv}) < H$.

If a voice session arrives due to a hand-off when a cell has C active sessions (either of voice or data) and H data sessions are in suspension, an arriving voice session make an active data session be terminated. Suppose an active data session on a m-type platform is terminated due to an arrival of a voice session on a g-type platform, then state variables are related by

$$\begin{aligned} &v(x_{hv}, g) = v(s, g) - 1 \\ &v(x_{hv}, i) = v(s, i), \quad i \neq g \\ &w(x_{hv}, m) = w(s, m) + 1 \\ &w(x_{hv}, i) = w(s, i), \quad i \neq m \\ &r(x_{hv}, i, j) = r(s, i, j) \end{aligned} \tag{100}$$

if $J(x_{hv}) = C$, $r(x_{hv}) = H$.

We let Λ_{hv} be the average rate at which hand-off arrivals of voice session impinge on the cell and F_{vg} denote the fraction of hand-off arrivals of voice session that are from g-type platforms. Initially we guess Λ_{hv} and F_{vg} but these values are determined by the dynamics of process itself using an iterative method [4], [10]. Then the transition flow into state s from x_{hv} due to hand-off arrival of a voice session is given by

$$\gamma_{hv}(s, x_{hv}, g) = \begin{cases} \Lambda_{hv} \times F_{vg} & J(x_{hv}) < C \\ Q_s(x_{hv}, g) \times \Lambda_{hv} \times F_{vg} & J(x_{hv}) = C, r(x_{hv}) < H \\ Q_t(x_{hv}, g) \times \Lambda_{hv} \times F_{vg} & J(x_{hv}) = C, r(x_{hv}) = H. \end{cases} \tag{101}$$

Hand-Off Arrival of Data Sessions

ACTIVE DATA SESSION HAND-OFF ARRIVALS

When the number of sessions, either of voice or data, is less than C in target cell, an arriving hand-off data session can obtain a connection. Thus a permissible state x_{hw0} is a predecessor state of s for a hand-off arrival of an active data session on a g-type platform when $J(x_{hw0}) < C$ and the state variables are related by

$$v(x_{hw0}, i) = v(s, i)$$
$$w(x_{hw0}, g) = w(s, g) - 1$$
$$w(x_{hw0}, i) = w(s, i), \quad i \neq g \qquad (102)$$
$$r(x_{hw0}, i, j) = r(s, i, j).$$

When an active data session attempts a hand-off, it will become a 1-suspended session if the target cell has C active sessions (either of voice or data) *but* the total number of suspended sessions in that cell is less than H. A transition into state s, due to a hand-off arrival of active data session on a g-type platform when the cell is in state x_{hw0}, will cause the state variable $r(x_{hw0}, g, 1)$ to increase by 1. Thus a permissible state x_{hw0} is a predecessor state of s for a hand-off arrival of an active data session on a g-type platforms when $J(x_{hw0}) = C$ and $r(x_{hw0}) < H$, if the state variables are related by

$$v(x_{hw0}, i) = v(s, i)$$
$$w(x_{hw0}, i) = w(s, i)$$
$$r(x_{hw0}, g, 1) = r(s, g, 1) - 1 \qquad (103)$$
$$r(x_{hw0}, g, j) = r(s, g, j), \quad j \neq 1$$
$$r(x_{hw0}, i, j) = r(s, i, j), \quad i \neq g.$$

We let Λ_{hw} be the average rate at which hand-off arrivals of active data session impinge on the cell and F_{wg} denote the fraction of hand-off arrival of active data session that are from g-type platforms. Also, we initially guess Λ_{hw} and F_{wg} but these values are determined by the dynamics of process. The corresponding transition flow for active data session hand-off attempts is given by

$$\gamma_{hw0}(s, x_{hw0}, g) = \Lambda_{hw} \times F_{wg}. \qquad (104)$$

K-SUSPENDED SESSION HAND-OFF ARRIVALS

When the number of sessions, either of voice or data, is less than C in target cell, a hand-off attempt of k-suspended session will be succeed. Thus a permissible state x_{hwk} is a predecessor state of s for hand-off arrival of a k-suspended session on a g-type platform when $J(x_{hwk}) < C$ and the state variables are related by

$$\begin{aligned} &v(x_{hwk}, i) = v(s, i) \\ &w(x_{hwk}, g) = w(s, g) - 1 \\ &w(x_{hwk}, i) = w(s, i), \quad i \neq g \\ &r(x_{hwk}, i, j) = r(s, i, j). \end{aligned} \tag{105}$$

If, 1) there is no channel available in the target cell; but, 2) the total number of suspended sessions in the cell is less than the maximum number of supportable suspended sessions, H; and, 3) the counter of the mobile is less than $N+1$; then a hand-off attempt will fail and a k-suspended session will become a $(k+1)$-suspended session in the target cell. Thus a permissible state x_{hwk} is a predecessor state of s for a hand-off arrival of a k-suspended session on a g-type platform when $J(x_{hwk}) = C$ and $r(x_{hwk}) < H$, if the state variables are related by

$$\begin{aligned} &v(x_{hwk}, i) = v(s, i) \\ &w(x_{hwk}, i) = w(s, i) \\ &r(x_{hwk}, g, k+1) = r(s, g, k+1) - 1 \\ &r(x_{hwk}, g, j) = r(s, g, j), \quad j \neq (k+1) \\ &r(x_{hwk}, i, j) = r(s, i, j), \quad i \neq g \end{aligned} \tag{106}$$

It should be recalled that when the counter of the terminal indicates $N+1$, then the terminal has attempted the maximum allowable number of reconnection attempts. This data session will be forced into termination. So, a failure of a hand-off attempt of a data session when the counter of the terminal is N causes the session to be forced into termination.

Let $\Lambda_r(k)$ be the average rate at which hand-off arrivals of k-suspended session impinge on the cell and let $F_{rg}(k)$ denote the fraction of hand-off arrivals of k-suspended session that are from g-type platforms. We initially guess $\Lambda_r(k)$ and $F_{rg}(k)$ but those values are, also, determined by the dynamics of process. Then the corresponding transition flow is given by

$$\gamma_{hwk}(s, x_{hwk}, g) = \Lambda_r(k) \times F_{rg}(k). \tag{107}$$

Hand-Off Departure of Voice Sessions

A transition into state s, due to a hand-off departure of a voice session on a g-type platform when the cell is in state x_{dv}, will cause the state variable $v(x_{dv}, g)$ to be decreased by 1. Thus a permissible state x_{dv} is a predecessor state of s for completion of a voice session on a g-type platform if the state variables are related by

$$\begin{aligned} &v(x_{dv}, g) = v(s, g) + 1 \\ &v(x_{dv}, i) = v(s, i), \quad i \neq g \\ &w(x_{dv}, i) = w(s, i) \\ &r(x_{dv}, i, j) = r(s, i, j). \end{aligned} \tag{108}$$

The transition flow into state s from x_{dv} due to a hand-off departure of a voice session is given by

$$\gamma_{dcv}(s, x_{dv}, g) = \mu_D(g) \times v(x_{dv}, g). \tag{109}$$

Hand-off Departure of Data Sessions

ACTIVE DATA SESSION HAND-OFF DEPARTURES

A transition into state s, due to a hand-off departure of an active data session on a g-type platform when the cell is in state x_{dw0}, will cause the state variable $w(x_{dw0}, g)$ to be decreased by 1. Thus a permissible state x_{dw0} is a predecessor state of s for a hand-off departure of an active data session on a g-type platform, if the state variables are related by

$$\begin{aligned} &v(x_{dw0}, i) = v(s, i) \\ &w(x_{dw0}, g) = w(s, g) + 1 \\ &w(x_{dw0}, i) = w(s, i), \quad i \neq g \\ &r(x_{dw0}, i, j) = r(s, i, j). \end{aligned} \tag{110}$$

The corresponding transition flow is given by

$$\gamma_{dw0}(s, x_{dw0}, g) = \mu_D(g) \times w(x_{dw0}, g). \tag{111}$$

K-SUSPENDED SESSION HAND-OFF DEPARTURES

A transition into state s, due to a hand-off departure of a k-suspended session on a g-type platform when the cell is state x_{dwk}, will cause the state variable $r(x_{dwk}, g, k)$ to be deceased by

1. Thus a permissible state x_{dwk} is a predecessor state of s for a hand-off departure of a k-suspended session on a g-type platform, if the state variables are related by

$$\begin{aligned} & v(x_{dwk}, i) = v(s,i) \\ & w(x_{dwk}, i) = w(s,i) \\ & r(x_{dwk}, g, k) = r(s, g, k) + 1 \\ & r(x_{dwk}, g, j) = r(s, g, j), \quad j \neq k \\ & r(x_{dwk}., i, j) = r(s, i, j), \quad i \neq g. \end{aligned} \tag{112}$$

The corresponding transition flow is given by

$$\gamma_{dwk}(s, x_{dwk}, g) = \mu_D(g) \times r(x_{dwk}, g, k). \tag{113}$$

Retry Attempt

If the cell has fewer than C active sessions when a terminal makes a reconnection attempt (either of hand-off or of retry) for a k-suspended session, it will simply succeed and the session will be active again. A transition into state s, due to a retry for k-suspended session on a g-type platform, when $J(x_r) < C$ and the cell state x_r, will cause the state variable $v(x_r, g)$ to be increased by 1 and $r(x_r, g, k)$ to be decreased by 1. Thus, a permissible state x_r is a predecessor state of s for a retry attempt of a k-suspended session on a g-type platform, if $J(x_r) < C$ and the state variables are related by

$$\begin{aligned} & v(x_r, i) = v(s,i) \\ & w(x_r, g) = w(s, g) - 1 \\ & w(x_r, i) = w(s,i), \quad i \neq g \\ & r(x_r, g, k) = r(s, g, k) + 1 \\ & r(x_r, g, l) = r(s, g, l), \quad l \neq k \\ & r(x_r, i, j) = r(s, i, j), \quad i \neq g. \end{aligned} \tag{114}$$

If there are C sessions, either of voice or data, in a cell when the terminal makes a retry attempt for a k-suspended session, it will fail. When the counter in terminal indicates less than N, the terminal will wait for the next reconnection attempt. A transition into state s, due to failure of a retry attempt for a k-suspended session on a g-type platform, if the counter of the mobile indicates less than the maximum allowable number of reconnection attempts ($k<N$), and

the cell is in state x_r, will cause the state variable $r(x_r, g, k)$ to be decreased by 1 and $r(x_r, g, k+1)$ to be increased by 1. Thus a permissible state x_r is a predecessor state of s for a failure of a k-repeated trial on a g-type platform when $J(x_r) = C$ and $k \neq N$, if the state variables are related by

$$\begin{aligned} &v(x_r, i) = v(s, i) \\ &w(x_r, i) = w(s, i) \\ &r(x_r, g, k) = r(s, g, k) + 1 \\ &r(x_r, g, k+1) = r(s, g, k+1) - 1 \\ &r(x_r, g, j) = r(s, g, j), \qquad j \neq k \vee k+1 \\ &r(x_r, i, j) = r(s, i, j), \qquad i \neq g. \end{aligned} \tag{115}$$

When the counter of terminal indicates $N+1$ after failure of a retry attempt, then the terminal has attempted maximum allowable reconnection attempts. This session will be forced into termination. Thus a permissible state x_r is a predecessor state of s for the failure of N-reconnection attempt on a g-type platform when $J(x_r) = C$, if the state variables are related by

$$\begin{aligned} &v(x_r, i) = v(s, i) \\ &w(x_r, i) = w(s, i) \\ &r(x_r, g, N) = r(s, g, N) + 1 \\ &r(x_r, g, j) = r(s, g, j), \qquad j \neq N \\ &r(x_r, i, j) = r(s, i, j), \qquad i \neq g. \end{aligned} \tag{116}$$

The corresponding transition flow is given by

$$\gamma_r(s, x_r, g, k) = \mu_r(k, g) \times r(x_r, g, k). \tag{117}$$

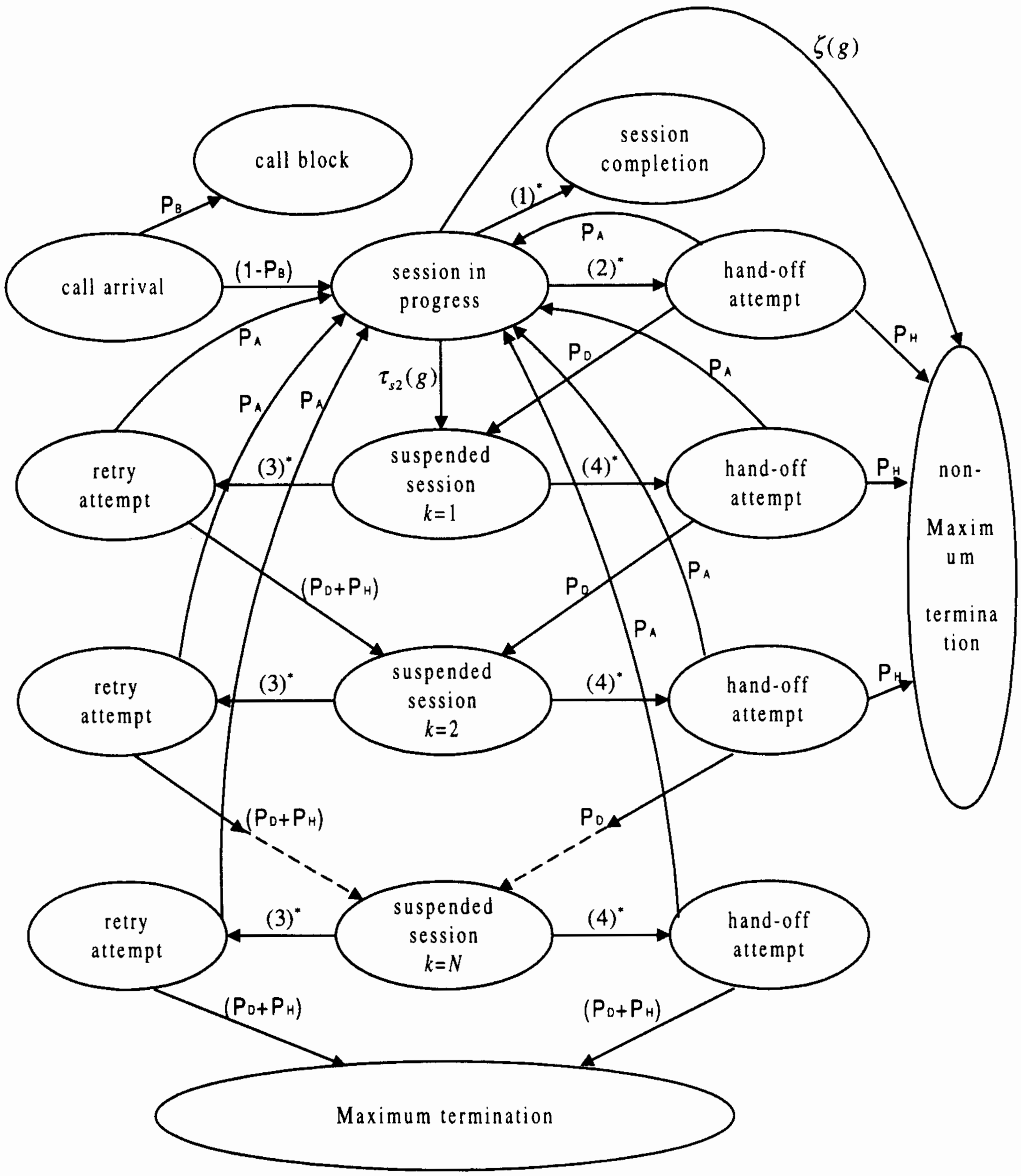

Figure 1: Flow graph of event of data type session, where $(1)^* = \dfrac{\mu_w(g)}{\mu_D(g)+\mu_w(g)}$, $(2)^* = \dfrac{\mu_D(g)}{\mu_D(g)+\mu_w(g)}$, $(3)^* = \dfrac{\mu_r(k,g)}{\mu_D(g)+\mu_r(k,g)}$, and $(4)^* = \dfrac{\mu_D(g)}{\mu_D(g)+\mu_r(k,g)}$.

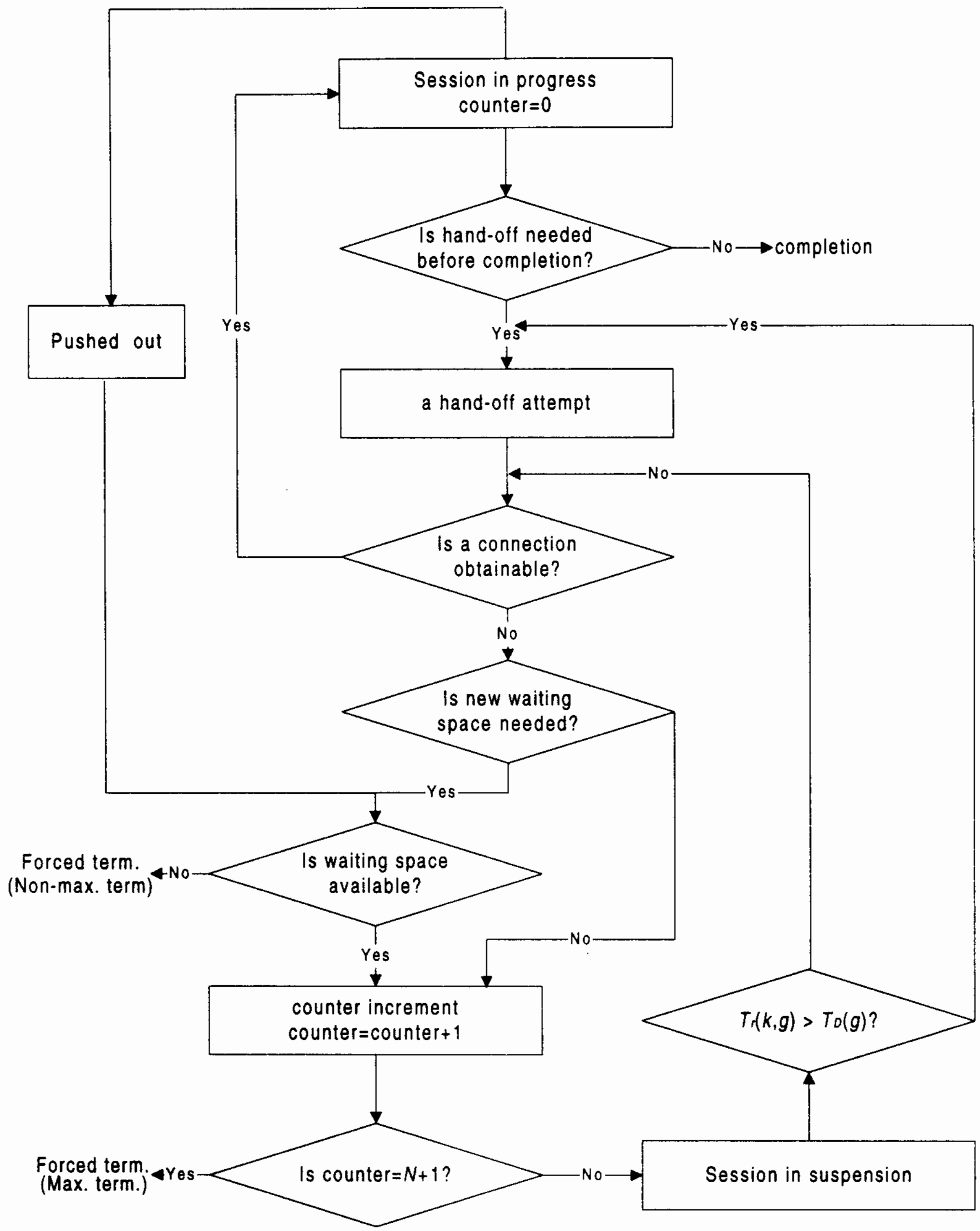

Figure 2: Flow chart of operation of reconnection counter.

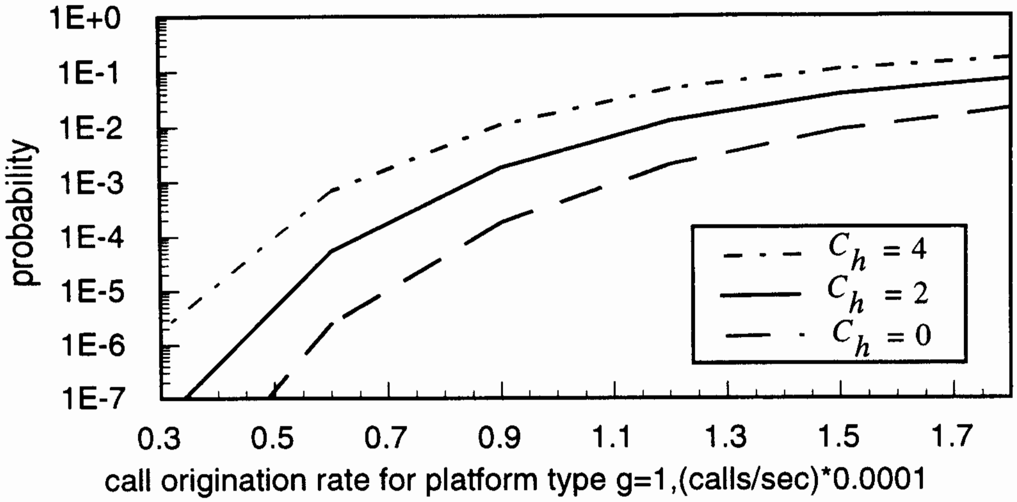

Figure 3: Blocking probability of voice type sessions: $C = 15$, $G = 2$, $\nu(1,0) = \nu(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100$s, $\overline{T}_w(1) = \overline{T}_w(2) = 20$s, $\overline{T}_D(1) = 500$s, $\overline{T}_D(2) = 100$s were assumed.

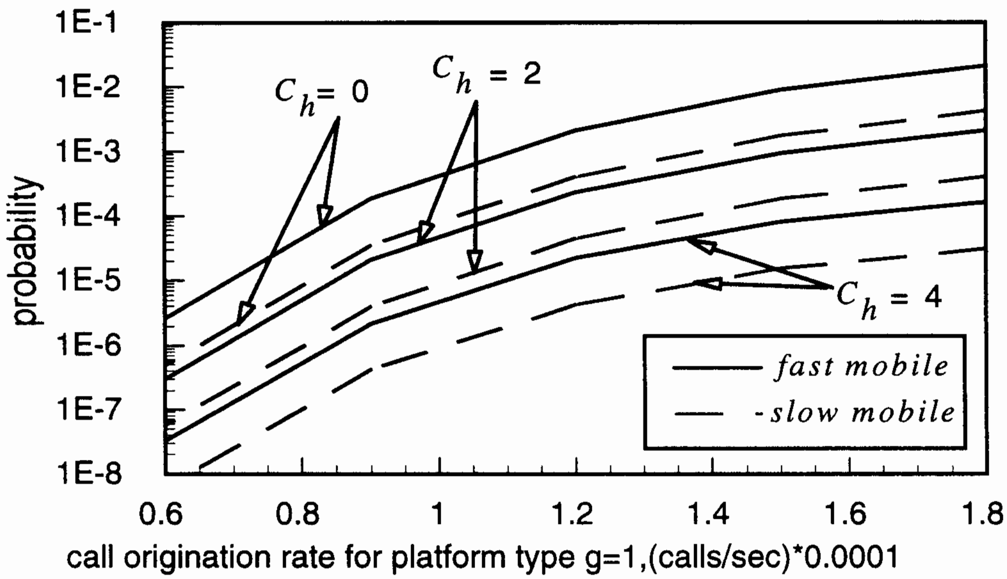

Figure 4: Forced termination probability of voice type sessions: $C = 15$, $G = 2$, $\nu(1,0) = \nu(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100$s, $\overline{T}_w(1) = \overline{T}_w(2) = 20$s, $\overline{T}_D(1) = 500$s, $\overline{T}_D(2) = 100$s were assumed.

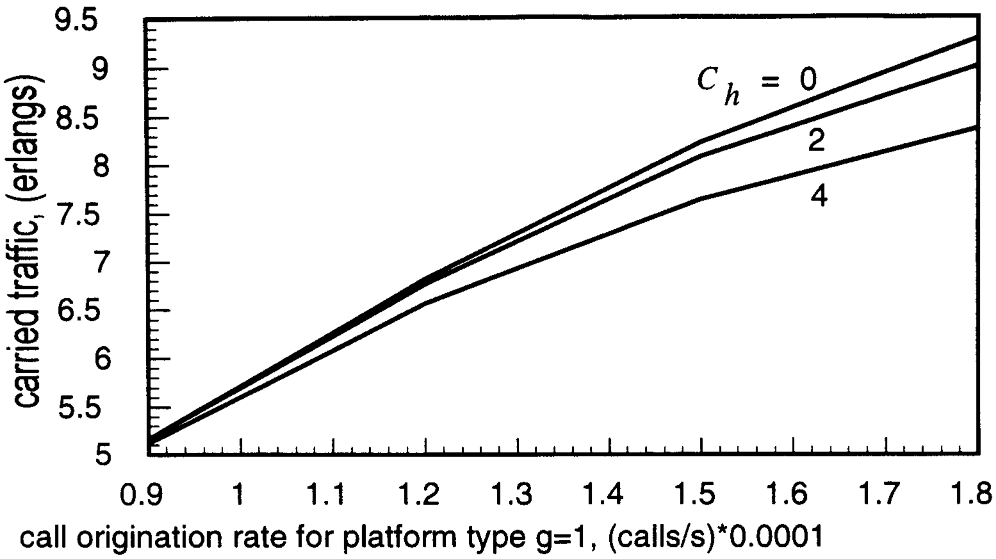

Figure 5: Carried traffic of voice type sessions: $C = 15$, $G = 2$, $v(1,0) = v(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100s$, $\overline{T}_w(1) = \overline{T}_w(2) = 20s$, $\overline{T}_D(1) = 500s$, $\overline{T}_D(2) = 100s$ were assumed.

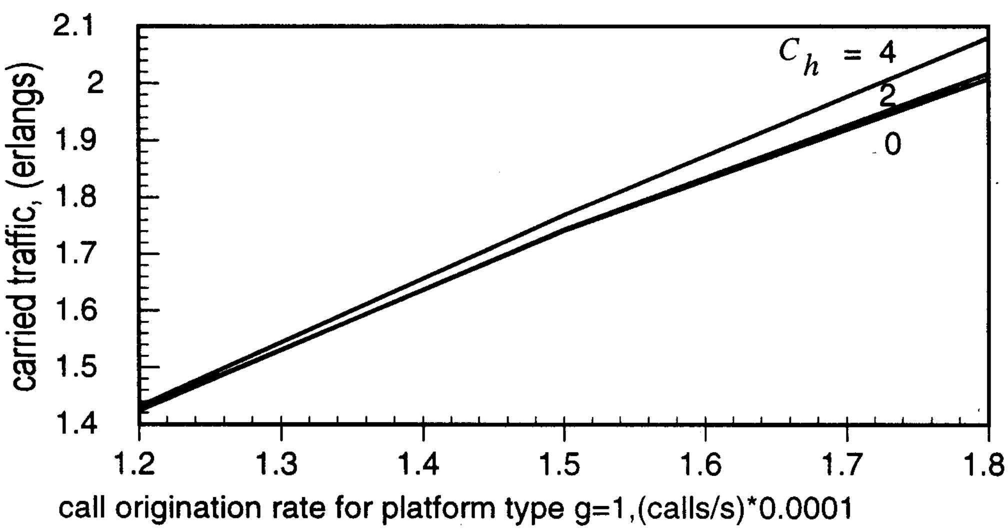

Figure 6: Average number of data sessions (active + suspended) per cell: $C = 15$, $G = 2$, $v(1,0) = v(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100s$, $\overline{T}_w(1) = \overline{T}_w(2) = 20s$, $\overline{T}_D(1) = 500s$, $\overline{T}_D(2) = 100s$, $N = 3$, $H = 3$, $\overline{T}_r(k, g) = 10s$ were assumed.

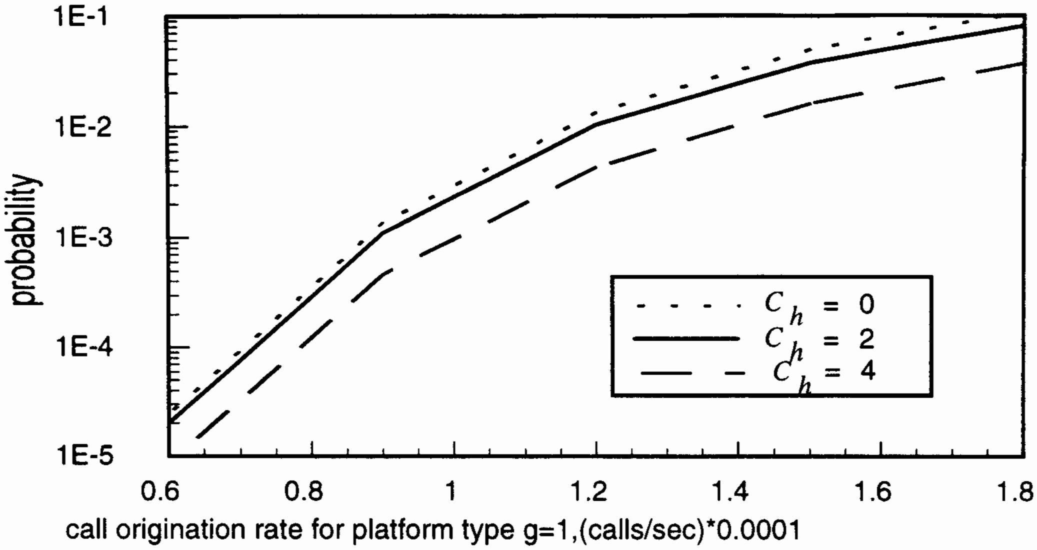

Figure 7: Blocking probability of data type sessions: $C = 15$, $G = 2$, $\nu(1,0) = \nu(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100s$, $\overline{T}_w(1) = \overline{T}_w(2) = 20s$, $\overline{T}_D(1) = 500s$, $\overline{T}_D(2) = 100s$, $N = 3$, $H = 3$, $\overline{T}_r(k, g) = 10s$ were assumed.

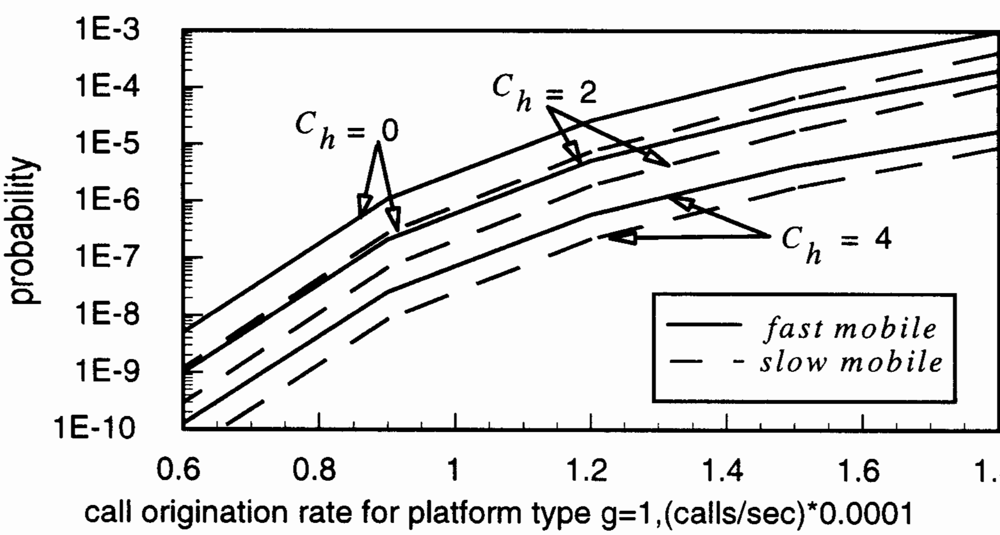

Figure 8: Forced termination probability of data type sessions: $C = 15$, $G = 2$, $\nu(1,0) = \nu(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100s$, $\overline{T}_w(1) = \overline{T}_w(2) = 20s$, $\overline{T}_D(1) = 500s$, $\overline{T}_D(2) = 100s$, $N = 3$, $H = 3$, $\overline{T}_r(k, g) = 10$ s were assumed.

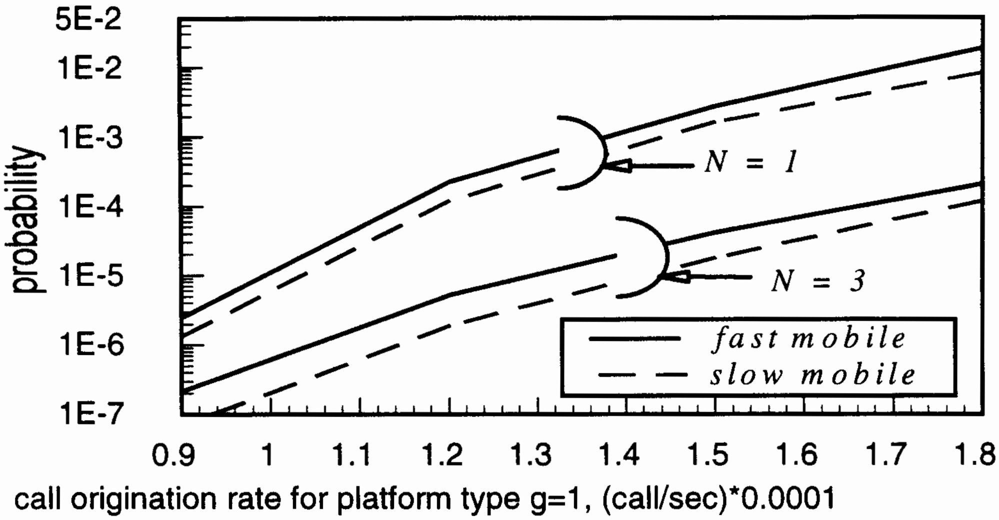

Figure 9: Forced termination probability of data type sessions: $C = 15$, $G = 2$, $v(1,0) = v(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100$s, $\overline{T}_w(1) = \overline{T}_w(2) = 20$s, $\overline{T}_D(1) = 500$s, $\overline{T}_D(2) = 100$s, $C_h = 2$, $H = 3$, $\overline{T}_r(k, g) = 10$ s were assumed.

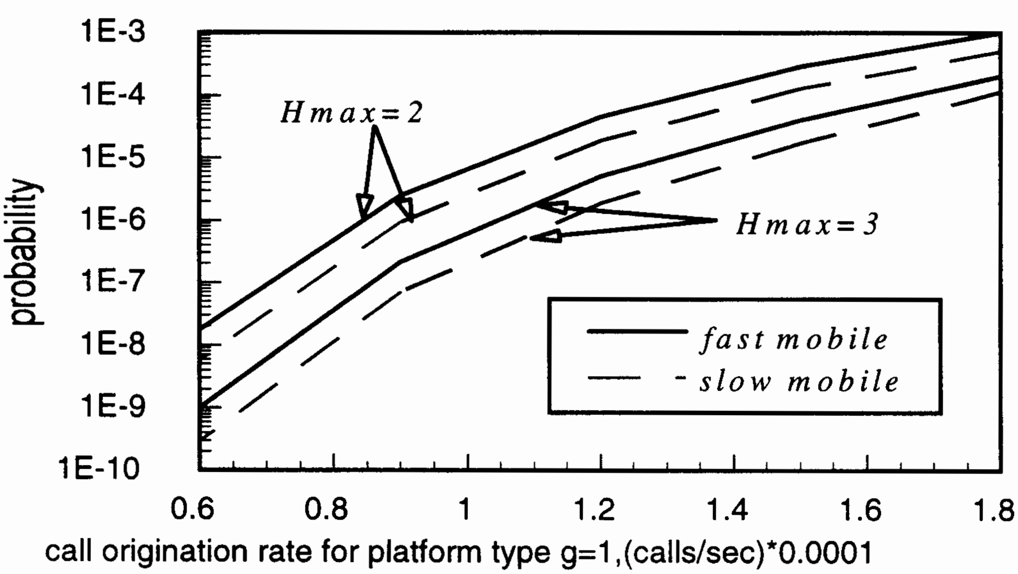

Figure 10: Forced termination probability of data type sessions: $C = 15$, $G = 2$, $v(1,0) = v(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100$s, $\overline{T}_w(1) = \overline{T}_w(2) = 20$s, $\overline{T}_D(1) = 500$s, $\overline{T}_D(2) = 100$s, $C_h = 2$, $N = 3$, $\overline{T}_r(k, g) = 10$ s were assumed.

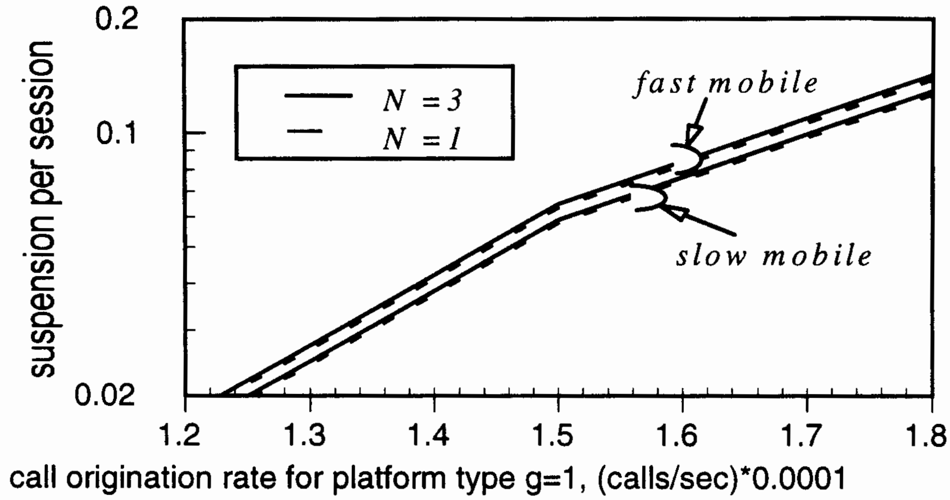

Figure 11: Average number of suspensions per session: $C = 15$, $G = 2$, $v(1,0) = v(2,0) = 300$, $\alpha(2) = \Lambda_{nv}(2) / \Lambda_{nv}(1) = 1.0$, $\beta(1) = \Lambda_{nw}(1) / \Lambda_{nv}(1) = 1.0$, $\beta(2) = \Lambda_{nw}(2) / \Lambda_{nv}(1) = 1.0$, $\overline{T}_v(1) = \overline{T}_v(2) = 100$s, $\overline{T}_w(1) = \overline{T}_w(2) = 20$s, $\overline{T}_D(1) = 500$s, $\overline{T}_D(2) = 100$s, $C_h = 2$, $H = 3$, $\overline{T}_r(k, g) = 10$ s were assumed.