Performance Modeling of Level Switching in Multitier Mobile Communication Systems

C.K. Ng and H.W. Chan

Abstract—The multitier environment of mobile communication systems is an integration of different mobile systems such as cellular, cordless, and satellite systems for mobile users to communicate with anyone, anywhere, and anytime. An effective and efficient scheme is needed for mobile terminals to dynamically select the most appropriate tier of mobile system for optimizing communication cost and performance. This paper presents a model for studying the effect of level switching in a multitier environment and for the performance evaluation of level switching schemes. This model is based on the continuous-time stochastic process, which is parameterized by a set of state-transition rates and costs derived from the characteristics of the mobile systems and terminals. Several level switching schemes based on different triggering events are proposed and the performance is analyzed using the proposed model. Level switching is found to be an effective way to reduce communication costs. The performance gain increases with the frequency of the switching activities and the level switching costs are insignificant in most cases. By using the model, the most cost-effective switching scheme can be selected to match the characteristics of individual mobile users and mobile systems for optimizing overall system performance.

Index Terms—Mobile communication systems, performance of systems, modeling techniques.

◆

1 INTRODUCTION

NOWADAYS, different types of *mobile communication systems* (MCSs) have been established to provide communication services to *mobile users* (MUs). These include cordless systems, cellular systems, and mobile satellite systems. Each MCS has specific characteristics designed for particular groups of MUs. For example, cordless telephone systems are developed to provide mobility support within residential areas, whereas mobile satellite systems are designed for areas without any wired network or other wireless systems. Each MCS has its own benefits and deficiencies and MUs need to carefully select the MCS that best fits their mobile communication requirements. However, many MUs may require services from more than one MCS. For example, they need cellular systems when they go outdoors and they prefer cordless systems when they are indoors. They may require mobile satellite systems when they go to areas like mountains and oceans that are not being served by cellular systems. In these cases, the MUs need to subscribe to the services of multiple MCSs. They may need to carry a *mobile terminal* (MT) for each MCS. Even if a single MT can support all the MCSs, they still need to manually switch to the systems they want.

A possible solution to this problem is to establish a new MCS to replace all existing MCSs in providing worldwide services to different kinds of MUs. However, it is very ineffective, costly, and time-consuming. Even if we can afford to build such an MCS, when the MUs demand more

bandwidth and services (e.g., multimedia and Internet access), the MCS has to be constantly upgraded to meet the increasing requirements. The development of new MCSs is usually to complement services provided by existing MCSs rather than to replace them, so the easiest and fastest way to fulfill the requirements of most MUs is to integrate various MCSs to form a multitier environment. The evolution of wireless communications is toward this direction in defining a common standard worldwide to unify the diverse systems into a seamless radio infrastructure capable of offering a wide range of services in many different radio environments [1]. Roaming service currently provided by most cellular systems is a good starting point for the integration and MUs are expecting the integration to be more transparent so that they can treat all the MCSs they subscribe to as a single system.

The *multitier environment of mobile communication systems* (multitier system) is an integration of different MCSs with different characteristics that complement each other. It provides MUs with the ability to communicate with anyone, anywhere, and anytime through integration of coverage area, radio frequency bandwidth, and services. Other advantages of the multitier system include the reduction in *location management* (LM) cost by using the MCS with the optimal cell sizes to track the MUs' locations [2], [3] and the reduction in *conversation cost* (CV cost) by selecting the MCS with the lowest CV cost.

The concept of multitier cellular systems or hierarchical cell structure has been described in various papers and articles on the third-generation mobile system. The Universal Mobile Telecommunications System (UMTS) should have integration in the fixed network for any MTs to have access to the picocell, microcell, macrocell, and satellite networks [4]. The International Telecommunication Union

---

- *The authors are with the Department of Computer Science, The University of Hong Kong, Pokfulam, Hong Kong. E-mail: {ckng, hwchan}@cs.hku.hk.*

(ITU) has specified architecture and system standards for the International Mobile Telecommunications System 2000 (IMT 2000) that provides a structural basis for the properties of a global network [5], [6]. These specifications include a hierarchical cell structure providing heterogeneous network services that support radio environments ranging from high-capacity picocells to urban terrestrial microcells and macrocells and to large satellite cells. The procedures required for Personal Communications Services (PCS) in new communication systems such as UMTS and IMT 2000 are described in [7]. These include mobility, customization, service-profile management, communication, and security procedures. The next generation of wireless communication is based on a global system of fixed and wireless mobile services with the integration of heterogeneous services across network providers and network backbones as well as geographical regions [8].

There are several issues with the feasibility of implementing the multitier system including interoperability, quality of service (QoS), security, billing, etc. On the system *interoperability* issue in enabling MTs to switch between MCSs, MTs should have the ability to access each MCS through different air-interface standards and MCSs with different network infrastructures should be able to communicate with each other to hand over the call from one system to another. The former issue can be solved by using multistandard MTs and subscribing to the services of different MCSs by the MUs. The latter issue can be overcome by defining standards on the interfaces, signaling protocol, sharing of knowledge, definition and management of services, and distribution of service functions for MCSs with different infrastructures to comply with [7]. Horizontal and vertical handovers are required within the peer systems and across systems at lower and higher hierarchical levels [9]. The use of software radio technology is likely to play a major role in the development of multistandard terminals with the capability of adapting themselves to available systems for universal roaming. Even in IMT 2000, there are at least six standards for the air interface. Efforts toward harmonization among various standards have been made in using a layered approach to separate the radio interface functions into transmission-dependent and transmission-independent functions so as to facilitate interoperation among systems. Another approach to simplify the system-interface problem is by defining a mobile Internet architecture in which Internet Protocol (IP) is used for the delivery of both voice and multimedia data between different MCSs. A multitier wireless-communication architecture achieved by the integration of Mobile IP and cellular IP was proposed to provide multimedia services in addition to voice communication [10]. The multitier system can improve *QoS* by assigning necessary bandwidth from different MCSs to meet the demand of MUs. It also improves services by preventing call rejection and call dropout due to an MT moving out of the radio coverage of one MCS. Traffic and resource management algorithms were proposed to exploit the potential advantages of the multitier system by the

allocation of each connection to the most convenient hierarchical layer according to the user-mobility profile and traffic nature [11]. Channel-assignment and handoff-management techniques were proposed to provide traffic flow control between layers so as to maximize the system capacity without degrading the traffic QoS [12]. A handoff method was proposed to improve QoS and resource-switching management to reduce data-packet loss for mobile multimedia communication in cellular IP and Mobile IP networks [13]. The *security* issue in the multitier system involves the authentication of an MT to different MCSs and the authentication of an MCS to other MCSs for call handover. The former issue can be solved by registering the MT to each of the MCSs. For the latter issue, the MCSs can have mutual agreements for transferring calls among themselves. Authenticated network connections between these MCSs can then be established for the call transfer. The *billing* issue can be resolved by setting up a charging scheme according to the connection time or transmitted data volume and the charging rate of each of the MCSs. An MU can be billed by a single MCS on the total airtime used or total data volume transmitted in all serving MCSs. After the fee has been collected, the MCS can distribute the fee among other MCSs according to the charging rates in their business agreements.

As the multitier system is regarded as feasible and promising by the organizations mentioned above, many researchers are making efforts to solve the technical issues in different aspects. However, they usually concentrate on specific problems in a particular type of multitier system. For example, the optimal design of a multitier system to determine the cell sizes for each tier and the portion of channels assigned to each tier was considered in [2]. Mobility issues such as registration, LM, and call routing in cordless-cellular network integration were studied in [14]. The design of an optimized GSM-satellite integration system architecture was described in [15]. A multilayer scheme for hierarchical cellular structures to tackle the mobility management problem by using the mobility of MT to determine the registration level was presented in [16]. Macrocell/microcell selection schemes based on velocity-estimation methods in a multitier cellular system that use the sojourn time in a microcell overlapping region were proposed in [17], [18]. Algorithms for assigning components of multimedia calls to different tiers based on a predefined sequence according to the bandwidth requirements of the components were discussed in [19]. An enhanced data-traffic management policy that allocates each connection to the most convenient hierarchical layer based on the user-mobility profile and traffic characteristics was proposed in [20]. The registration traffic analysis and service availability problem for two-tier wireless networks was studied in [21].

However, a key success factor of the multitier system has not been addressed by most previous research. This is to have an effective and efficient scheme for MTs to dynamically select the most appropriate tier of MCS for both LM and conversation so as to optimize signaling

traffic and performance. In order to find the best one among various possible system-selection schemes, a performance model and the corresponding evaluation mechanism are needed to evaluate their performance under different conditions. In the performance analysis of MCSs, the stochastic model is widely used to model the activities of MTs. For example, the movement of MT was modeled by a continuous-time Markov process in [22], [23], [24]. A dynamic predictive LM scheme using a continuous-time Markovian mobility model characterized by cell-to-cell transition probabilities was presented in [22]. The concept of implicit location registration using a continuous-time Markov chain model for performance analysis was introduced in [23]. The cell identification codes were proposed in [24] for tracking MUs. The movements were assumed to be stochastic and independent from one MU to another in the performance analysis. The calls were assumed to be initiated by the MUs as a Poisson process. The processes by which the MU calls out and receives incoming calls were also assumed to be Poisson processes in [25], [26]. A mobility-tracking mechanism that combines a movement-based location-update policy with a selective paging scheme was introduced in [25]. Hierarchical multiresolution registration structure was proposed in [26] for mobility tracking. The call duration was also assumed exponentially distributed for the performance analysis of the distributed multiagent LM scheme in [27]. An analytical model for performance evaluation on channel allocation schemes in single-tier cellular communication networks was shown in [28]. However, there is a lack of a performance model for the multitier system.

In this paper, a model is developed to represent the multitier environment of MCSs as a single system with different levels of MCSs. A continuous-time stochastic model parameterized by a set of state-transition rates and costs is constructed. The state-transition rates are derived from characteristics of MCSs and behavior of individual MTs. The costs denote signaling traffic in communication activities like conversation, LM, and *level switching* (LS). In particular, the LS cost has not been considered by most previous works. The costs also denote the penalty of calls being rejected or dropped. This factor is important in providing services with quality but was usually neglected by previous works. Several LS schemes for optimizing total system cost are proposed based on triggering events, which distinguishes this paper from other research works that only consider which level to select. The performance of these LS schemes in terms of total system cost is analyzed and compared by using the stochastic model developed. From steady-state analysis, we can obtain the behavior of the LS schemes with respect to variations in the characteristics of MCSs and MUs. Also from the comparison, we can find the optimal LS scheme under any given conditions.

This paper is structured as follows: Section 2 describes the performance model in detail. Section 3 presents the proposed LS schemes. Section 4 compares and analyzes the performance of the LS schemes using the developed model.
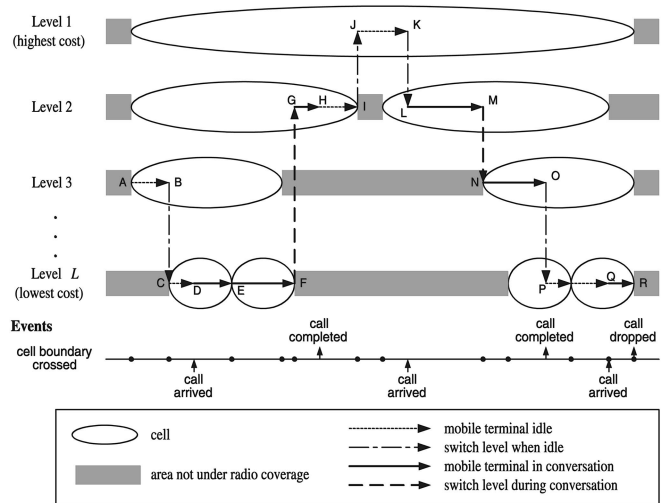

Fig. 1. Performance model structure.

Section 5 describes an enhanced performance model. Finally, a conclusion is given in Section 6.

2 PERFORMANCE MODEL

2.1 Structure

The performance model for the multitier system consists of four main entities: *mobile terminal*, *level*, *service area*, and *cell*. The MT is the device used by the MU to communicate with others through one or more MCSs. The MCSs subscribed by the MU are defined as *levels* of the multitier system, as shown in Fig. 1. These levels are arranged in descending order of *communication costs* (COMM costs) including CV, LM, and LS costs, which will be defined in Section 2.3. The *service area* of a level is the geographical area in which the MCS can provide services. The *probability of under coverage* is defined as the probability that the MTs can receive the radio signals from the MCS at their current locations. The service area of each level is divided into small regions called *cells* that overlap at the outer boundaries. Each cell has its own radio equipment for transmission to and reception from MTs within the cell. The cell size is determined by the dimensions of the geographic area served by the base station of the cell. It is designed to handle the expected number of MUs in the cell. Each cell is assigned a number of frequency bands called channels for communication. It is assumed that MTs can always find unoccupied channels to communicate with the MCS. Within a level, the size and COMM costs of all the cells are assumed to be the same. Each MT can only register in one level for conversation or LM at a time. When the MT has more than one level to select, it can switch from one level to another more favorable level.

A *mobile-system-centric* approach is usually employed in performance models in which each MCS tracks the activities of all MTs entering and leaving the system. This traditional approach may be good for a single system, but it is difficult to model the switching activities in a multitier system. We employ the *mobile-terminal-centric* approach in this model in which the activities of each MT navigating among all the
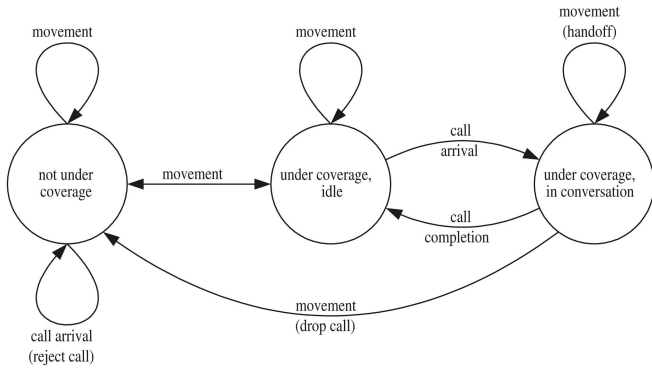
Fig. 2. State-transition diagram for a single-level system.

MCSs are considered. When we optimize the performance of each MT, the overall performance of all the MCSs can also be optimized. Two major types of MT activities are considered. One is related to the calling activities and the other is related to the movement of the MT. Other minor activities of the MT such as powering on and off are not considered. The *call-related activities* include call arrival and call completion, which are parameterized by *call-arrival rate* and *call-completion rate*, respectively. *Call arrival* is the event that an incoming call arrives at the system to communicate with the MU. *Call completion* is the event that the MU finishes the conversation and then hangs up the MT. These two rates are used to model the calling pattern of the MT and are independent of the level in which the MT registered. The status of the MT will be switched between *idle* and *in conversation* by these activities. The *movement-related activities* are parameterized by the MT's rate of movement in terms of *cell-boundary crossing rate*. This rate is used to model the mobility of the MT and is level-dependent. When the MT moves among the cells in a certain level, it will be in either one of the two statuses: *not under coverage* or *under coverage*. The status is determined by the *probability of under coverage* of the particular level. For each level, we assume that the area in which the MT may travel is divided into fictitious cells of equal size. This area includes the service area and the area outside the radio coverage of the MCS. So, some cells are under radio coverage and some are not. The purpose of doing this is to maintain a constant cell-boundary crossing rate and probability of under coverage inside and outside the service area. When call-related and movement-related activities are combined, we have three states for a single-level system, as shown in the state-transition diagram in Fig. 2. From the diagram, it can be observed that there are only three events to trigger a change of states: call arrival, call completion, and movement of the MT.

The structure of the model is shown in Fig. 1 with a one-dimension example of L levels illustrating how the MT navigates among various levels under different situations. At **A**, the MT moves into the radio coverage of the system and registers for LM in Level 3, which is the accessible level with the lowest cost. When it enters the coverage area of Level L at **B**, it can switch to Level L (at **C**) for lower LM cost or can remain in Level 3 to save the LS cost. Assuming that it has switched to Level L and travels along to **D**, an

incoming call arrives and it starts a conversation. When it crosses the cell boundary as it moves from one cell to another at **E**, a handoff occurs. When it moves out of the coverage of Level L at **F**, it has to find another level to continue the conversation. As it is under the coverage of Level 1 and 2 but with lower CV cost under Level 2, it switches to Level 2 at **G**. At **H**, it becomes idle when the conversation completes. When it moves out of the coverage of Level 2 at **I**, it finds that it is only under the coverage of Level 1 so it switches to Level 1 at **J** for the system to continue the tracking of its location. When it moves to **K**, an incoming call arrives. Since it is under the coverage of both Level 1 and 2, it can either stay at Level 1 or switch to Level 2 to start the conversation. Assuming it switches to Level 2 at **L** for lower CV cost, when it moves to **M** and finds out that it is also under coverage of Level 3, it selects to switch to Level 3 at **N** to continue the conversation with lower cost. When it completes the conversation at **O**, it switches to Level L at **P** for lower LM cost instead of staying in Level 3. After it has started a conversation at **Q** and moves out of the coverage of all levels at **R**, it has to drop the call in progress. Under such a condition of not under coverage of any levels, the MT is defined to be registered in Level 0.

2.2 Stochastic Model

There are three events in the proposed model: *call arrival*, *call completion*, and *cell-boundary crossing*. These events are assumed to be Poisson arrival processes with independent negative exponential distributions. With this assumption, the system is modeled by a continuous-time discrete-state stochastic process with states denoted by $S_{xyz_1z_2z_3\cdots z_i\cdots z_L}$ representing the status of the MT in the system, where

$$x = \text{conversation indicator}$$
$$= \begin{cases} 0 & \text{idle} \\ 1 & \text{in conversation,} \end{cases}$$
$$y = \text{registered level}$$
$$= \begin{cases} 0 & \text{if} \quad z_i = 0 \quad \forall i = 1 \ldots L \\ 1 \ldots L & \text{given} \quad z_y = 1, \end{cases}$$
$$z_i = \text{status in Level } i \ (i = 1 \ldots L)$$
$$= \begin{cases} 0 & \text{not under coverage} \\ 1 & \text{under coverage.} \end{cases}$$

For example, in a three-level system, S_{00000} means that the MT is not under coverage of any levels and S_{02011} means that the MT is idle ($x = 0$) and registered in Level 2 ($y = 2$), not under coverage of Level 1 ($z_1 = 0$) but under coverage of Level 2 and 3 ($z_2 = z_3 = 1$).

The model can be represented by the following parameters:

$$L = \text{number of levels,}$$
$$\delta = \text{mean time between calls,}$$
$$\tau = \text{average conversation time per call,} \qquad (1)$$
$$\sigma_i = \text{average cell residence time in Level } i,$$
$$\beta_i = \text{probability of under coverage in Level } i.$$

L describes the structure of the system. δ and τ model the calling pattern of the MT, while β_i describes the
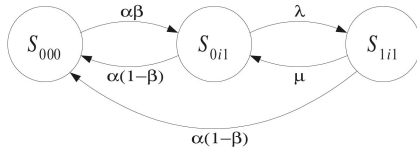
Fig. 3. Transition-rate diagram for single-level system (for level i, $\alpha = \alpha_i$ and $\beta = \beta_i$).

characteristics of the MCSs. σ_i is determined by the size of the cell in Level i and the speed of the MT. It models the mobility of the MT in Level i. In other words, L and β_i are independent of MTs, while δ, τ, and σ_i are usually different for individual MTs.

Let n be the number of states in the system and r_{jk} be the rate of transition from state j to state k ($j, k = 1 \cdots n$). The value of n depends on L and how the MT switches between levels. r_{jk} is given by one of the following rates according to the event triggering the change of states:

$$\lambda = \frac{1}{\delta} = \text{call arrival rate}, \tag{2}$$

$$\mu = \frac{1}{\tau} = \text{call completion rate}, \tag{3}$$

$$\alpha_i(1 - \beta_i) = \text{rate of moving out of coverage in Level } i, \tag{4}$$

$$\alpha_i \beta_i = \text{rate of moving under coverage in Level } i, \tag{5}$$

where

$$\alpha_i = \frac{1}{\sigma_i} = \text{cell boundary crossing rate in Level } i.$$

The first and second rates, (2) and (3), are for the call-arrival and call-completion events, respectively, while the last two rates, (4) and (5), are for the cell-boundary crossing event. For the call-arrival event, the state changes from $S_{0yz_1z_2\cdots z_i\cdots z_L}$ to $S_{1y'z_1z_2\cdots z_i\cdots z_L}$ and the rate is given by λ. x changes from 0 to 1, meaning that the MT starts the conversation. y changes to y', meaning that there may be a change in the registered level, that is, an LS may occur.

The rest of z_i remains unchanged. Similarly, for the call-completion event, the state changes from $S_{1yz_1z_2\cdots z_i\cdots z_L}$ to $S_{0y'z_1z_2\cdots z_i\cdots z_L}$ and the rate is given by μ. There are two situations for the cell-boundary crossing event. First, when the MT moves out of the coverage in Level i, the state changes from $S_{xyz_1z_2\cdots z_{i-1}1z_{i+1}\cdots z_L}$ to $S_{x'y'z_1z_2\cdots z_{i-1}0z_{i+1}\cdots z_L}$ and the rate is given by $\alpha_i(1 - \beta_i)$. Second, when the MT moves under the coverage in Level i, the state changes from $S_{xyz_1z_2\cdots z_{i-1}0z_{i+1}\cdots z_L}$ to $S_{xy'z_1z_2\cdots z_{i-1}1z_{i+1}\cdots z_L}$ and the rate is given by $\alpha_i\beta_i$. From the above four conditions of the transition events, we can derive the transition rates for the changes of states. The state-transition diagram for a single-level system (Level i) in Fig. 2 can be transformed to the transition-rate diagram in Fig. 3.

Steady-state analysis is used to find the steady-state probability π_j ($j = 1 \cdots n$) of the MT in a particular state j from the following set of balance equations [29], [30]:

sum of probabilities = unity

$$\Rightarrow \quad \sum_{j=1}^{n} \pi_j = 1 \tag{6}$$

transition rate into state j = transition rate out from state j

$$\Rightarrow \quad \sum_{k=1,\neq j}^{n} \pi_k r_{kj} = \pi_j \sum_{k=1,\neq j}^{n} r_{jk} \quad j = 1 \cdots n. \tag{7}$$

There are $n + 1$ equations and n unknowns. Any one of the equations in (7) is redundant and can be removed [30]. By solving these n linear equations, we can obtain the n steady-state probabilities.

2.3 Cost Model

Performance evaluation of the system is based on the total COMM cost in terms of cost units per unit of time. The costs of equipment setup and maintenance are not considered directly but are absorbed in various cost elements of conversation and LM. Fig. 4 shows the components of the COMM costs, the classification of the costs, and the
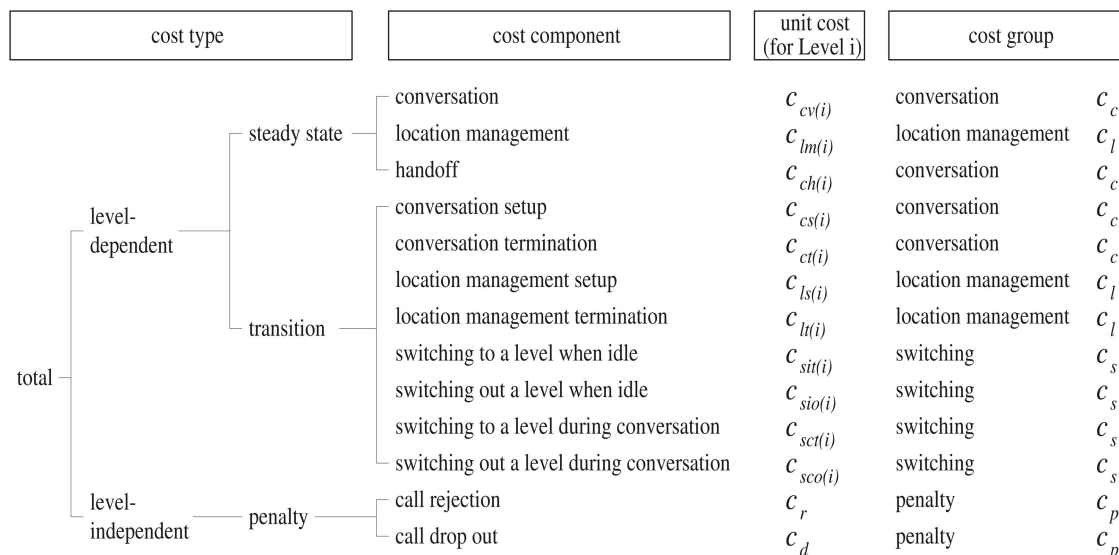

cost type				cost component	unit cost (for Level i)	cost group	
			steady state	conversation	$c_{cv(i)}$	conversation	c_c
				location management	$c_{lm(i)}$	location management	c_l
				handoff	$c_{ch(i)}$	conversation	c_c
	level-dependent			conversation setup	$c_{cs(i)}$	conversation	c_c
				conversation termination	$c_{ct(i)}$	conversation	c_c
				location management setup	$c_{ls(i)}$	location management	c_l
			transition	location management termination	$c_{lt(i)}$	location management	c_l
total				switching to a level when idle	$c_{sit(i)}$	switching	c_s
				switching out a level when idle	$c_{sio(i)}$	switching	c_s
				switching to a level during conversation	$c_{sct(i)}$	switching	c_s
				switching out a level during conversation	$c_{sco(i)}$	switching	c_s
	level-independent	penalty		call rejection	c_r	penalty	c_p
				call drop out	c_d	penalty	c_p

Fig. 4. Communication costs.

TABLE 1
Types of Transitions and Activities

Transition type	$T_a \in A_i$	$T_b \in B_j$	r_{ab}	d_{ab}
setup call	$x = 0$ $y = i > 0$ $z_i = 1$	$x' = 1$ $y' = j > 0$ $z'_j = 1$	λ	$c_{lt(i)} + c_{cs(j)}$
complete call	$x = 1$ $y = i > 0$ $z_i = 1$	$x' = 0$ $y' = j > 0$ $z'_j = 1$	μ	$c_{ct(i)} + c_{ls(j)}$
switch level when idle	$x = 0$ $y = i > 0$ $z_i = 1$	$x' = 0$ $y' = j \neq i$ $z'_j = 1$	$\alpha_j \beta_j$	$c_{sio(i)} + c_{sit(j)}$
switch level during conversation	$x = 1$ $y = i > 0$ $z_i = 1$	$x' = 1$ $y' = j \neq i$ $z'_j = 1$	$\alpha_j \beta_j$	$c_{sco(i)} + c_{sct(j)}$
move under coverage from out of coverage	$x = 0$ $y = 0$ $z_i = 0 \quad \forall i$	$x' = 0$ $y' = j > 0$ $z'_j = 1$	$\alpha_j \beta_j$	$c_{ls(j)}$
move out of coverage when idle	$x = 0$ $y = i > 0$ $z_i = 1$	$x' = 0$ $y' = 0$ $z'_j = 0 \quad \forall j$	$\alpha_i(1 - \beta_i)$	0
move out of coverage during conversation	$x = 1$ $y = i > 0$ $z_i = 1$	$x' = 0$ $y' = 0$ $z'_j = 0 \quad \forall j$	$\alpha_i(1 - \beta_i)$	$c_{ct(i)} + c_d$

Activity type	$T_a \in A_i$		r_a	d_a
conversation	$x = 1$ $y = i > 0$ $z_i = 1$		1	$c_{cv(i)}$
location management	$x = 0$ $y = i > 0$ $z_i = 1$		1	$c_{lm(i)}$
handoff	$x = 1$ $y = i > 0$ $z_i = 1$		$\alpha_i \beta_i$	$c_{ch(i)}$
call rejection	$x = 0$ $y = 0$ $z_i = 0 \quad \forall i$		λ	c_r

grouping of the costs for performance analysis. The COMM costs are classified under three types. The first type is the *steady-state cost*, which consists of the costs for conversation, LM, and handoff. The second type is the *transition cost*, which consists of the costs for setup and termination of conversation and of LM, switching on and off a level during conversation and when idle. The third type is the *penalty cost*, which consists of the penalties for call rejection and call dropout. *Call rejection* is the incident of rejecting an incoming call when the MT is idle. It is due to the MT not being under coverage of any levels. *Call dropout* is the incident that the conversation in progress is forced to terminate when the MT moves out of the coverage of all levels. The unit steady-state costs and unit transition costs are usually measured by the signaling traffic generated in the wireless and fixed networks by the MT. They depend on the level in which the MT registered, while the unit penalty costs are independent of the levels. As the levels are arranged in descending order of COMM costs, it is assumed that all level-dependent cost components are also in descending order. The unit LM cost, $c_{lm(i)}$, is not constant for all MTs since it depends on the mobility and calling

patterns of each MT and the LM strategy being used. Since only the steady-state behavior of a particular MT is considered, we can use the average value of $c_{lm(i)}$ calculated from δ, τ, and σ_i of the MT. Also, each MT can be assigned with unit penalty costs different from other MTs. Let

$$T_a = \text{state before transition}$$
$$= S_{xyz_1 z_2 z_3 \cdots z_L} \qquad a = 1 \cdots n,$$
$$T_b = \text{state after transition}$$
$$= S_{x'y'z'_1 z'_2 z'_3 \cdots z'_L} \qquad b = 1 \cdots n,$$
$$A_i = \text{set of states registered in Level } i \text{ before transition}$$
$$= \{S_{xiz_1 z_2 z_3 \cdots z_L}\} \qquad i = 0 \cdots L,$$
$$B_j = \text{set of states registered in Level } j \text{ after transition}$$
$$= \{S_{x'jz'_1 z'_2 z'_3 \cdots z'_L}\} \qquad j = 0 \cdots L,$$
$$\pi_a = \text{steady-state probability of } T_a,$$
$$r_a = \text{rate of activity happening in } T_a,$$
$$r_{ab} = \text{rate of transition from } T_a \text{ to } T_b,$$
$$d_a = \text{cost involved for the activity in } T_a,$$
$$d_{ab} = \text{cost involved in transition from } T_a \text{ to } T_b.$$

TABLE 2
Level Switching Schemes and Triggering Events

Perform level switching? / Event / Level Switching Scheme		Call arrival and completion (call-related)	Cell boundary crossing (movement-related)	
			under coverage of current level	out of coverage of current level
No Level Switching	NLS	No	No	No
Aggressive Level Switching	ALS	Not needed	Yes	Yes
Reactive Level Switching	RLS	No	No	Yes
Selective Level Switching	SLS	Yes	No	Yes

Note: Level switching is not needed in ALS for call-related events since the mobile terminal is always in the level with the lowest cost.

With the unit cost of each cost component defined in Fig. 4, the transition cost (C_X) and activity cost (C_A) for each type of transition or activity in Table 1 are calculated by

$$C_X = \sum_{i=0}^{L} \sum_{T_a \in A_i} \pi_a \sum_{j=0}^{L} \sum_{T_b \in B_j} r_{ab} d_{ab},$$

$$C_A = \sum_{i=0}^{L} \sum_{T_a \in A_i} \pi_a r_a d_a.$$

The total cost of the system (C_T) is given by the sum of the costs for all types of transitions and activities, that is,

$$C_T = \sum_{\text{all transitions}} C_X + \sum_{\text{all activities}} C_A. \qquad (8)$$

3   LEVEL SWITCHING SCHEMES

This section describes the ways of selecting levels for LM or conversation. The LS schemes are the criteria for the MT to determine *when* to switch from one level to another. The target level is predetermined to be the level that can provide the required resources (i.e., the radio coverage) with the lowest level-dependent costs. The target level can be determined easily since all levels are arranged in descending order of COMM cost, i.e., cost of Level i > cost of Level j if $i < j$. In other words, Level 1 has the highest cost and Level L has the lowest.

The major benefit of the LS schemes is reduction in the COMM costs. The level-dependent costs are reduced by selecting the level with the lowest cost among all the possible ones. The probabilities of rejecting or dropping calls are also reduced by switching to a level providing coverage for the communication. The call rejection and dropout penalty costs are subsequently reduced. However, the drawback is that additional costs are introduced in LS.

Four LS schemes are proposed based on the two types of MT activities: call-related and movement-related. The possible events for triggering LS and the corresponding schemes are shown in Table 2.

3.1   No Level Switching (NLS)

In the NLS scheme, the MT stays in one level only. It is the case when the MU subscribes to one MCS only. The MU will select the level with the lowest total cost statically according to its mobility and calling pattern. The MT will not switch to another level when it is not under coverage of the current level. The chances of rejecting incoming calls
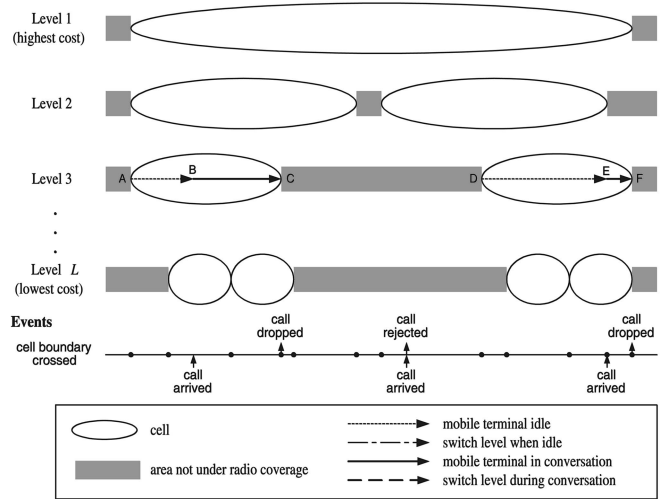

Fig. 5. Structure of the NLS scheme.

and dropping of calls in conversation are expected to be high. However, there are no LS costs involved. Strictly speaking, the NLS scheme may not be classified as a real LS scheme since there is actually no dynamic LS involved. However, it is used in the model as the base for the performance comparison of other LS schemes.

The structure of the NLS scheme for a system with four levels is shown in Fig. 5, where the MT only registers in Level 3. The conversation in progress will be dropped when the MT moves out of coverage at C and F. The incoming call between C and D will be rejected since there is no radio coverage in Level 3.

3.2   Aggressive Level Switching (ALS)

In the ALS scheme, the MT always tries to stay in the level with the lowest steady-state cost. It may perform frequent LS in order to switch to the level with the lowest cost, especially when it crosses the cell boundaries of various levels. The total CV and LM costs will be minimal but significant LS cost may result. As the MT will switch to another level when it is not under coverage of the current level, the chances of being unable to receive incoming calls and dropping of calls in conversation are expected to be low.

The structure of the ALS scheme for a system with four levels is shown in Fig. 6. The MT will switch levels from B to C, K to L, N to O, and P to Q when it enters cells of levels with lower costs. It will also switch levels from F to G to continue the conversation in progress and from I to J to continue the LM when it moves out of the coverage of the current level.

3.3   Reactive Level Switching (RLS)

In the RLS scheme, the MT remains in the current level whenever possible to reduce the number of LS. It switches to another level (the level with the required resources at the lowest cost) only when it moves out of the coverage of the current level. This will keep the chances of rejecting incoming calls and dropping of calls in conversation a minimum, just like the ALS scheme. The total LS cost will be reduced to minimal, but the total CV and LM costs will be higher than those of the ALS scheme.
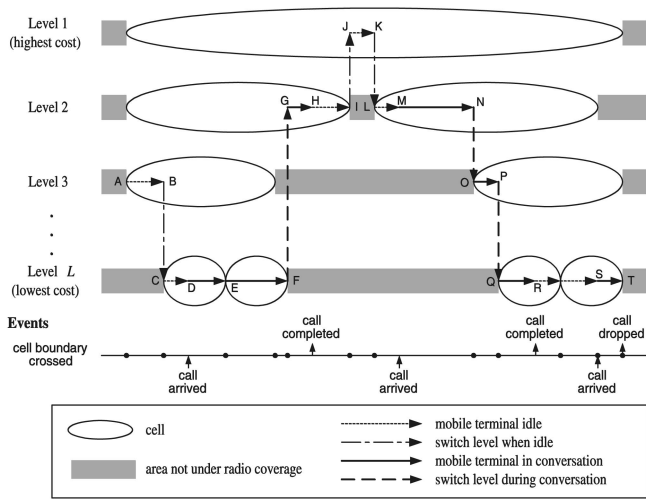
Fig. 6. Structure of the ALS scheme.

Fig. 8. Structure of the SLS scheme.

The structure of the RLS scheme for a system with four levels is shown in Fig. 7. The MT will switch levels from C to D and E to F when the current level cannot provide coverage for conversation, and from H to I when the current level cannot provide coverage for LM.

3.4 Selective Level Switching (SLS)

In addition to the events for triggering LS to prevent call rejection and call dropout in the RLS scheme, the MT will also perform LS for some selected events, such as call arrival and completion, in the SLS scheme. The MT will selectively perform LS to minimize the CV cost or LM cost. The total CV and LM costs will be reduced where the total LS cost will be higher than that of the RLS scheme.

The structure of the SLS scheme for a system with four levels is shown in Fig. 8. The MT will switch to a level with lower cost from B to C and from J to K to start a conversation when an incoming call arrives, and from L to M for lower LM cost when the conversation completes.
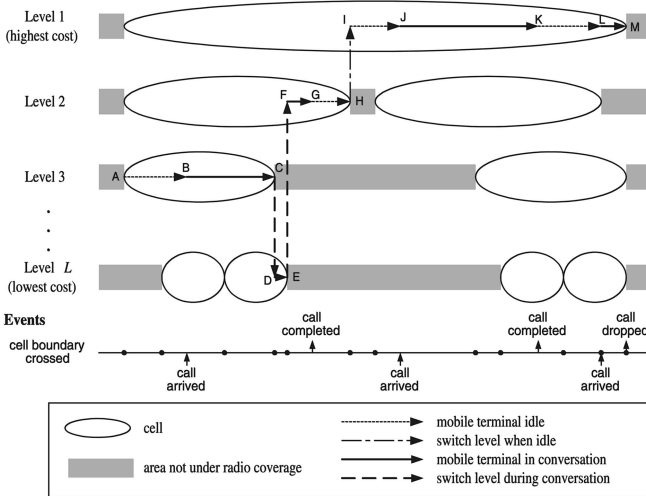

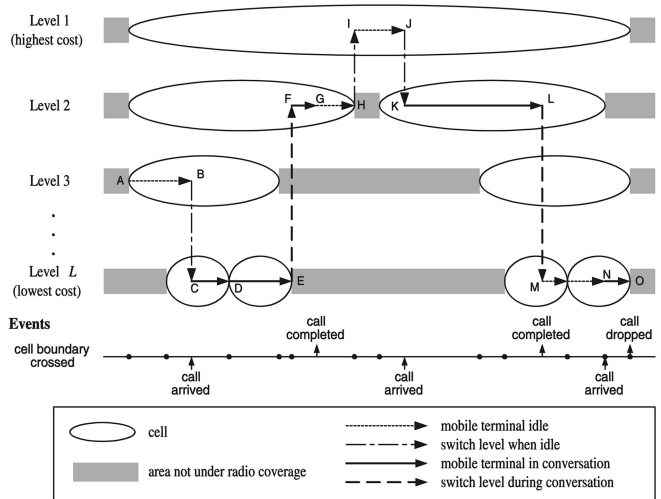Fig. 7. Structure of the RLS scheme.

4 PERFORMANCE ANALYSIS

4.1 Approach

The performance comparison of different LS schemes is based on the total COMM cost of the system employing the schemes in steady state. The scheme with the lowest total cost is considered to have the best performance. The performance analysis involves the following steps:

First of all, the parameters in (1) and the unit costs in Fig. 4 have to be specified for the model. The total number of states for each scheme can then be calculated by (9), whose derivations can be found in [31].

$$
\begin{aligned}
n_{\mathrm{NLS}} &= 3L, \\
n_{\mathrm{ALS}} &= 2^{L+1} - 1, \\
n_{\mathrm{RLS}} &= n_{\mathrm{SLS}} = 1 + L(2^L).
\end{aligned}
\tag{9}
$$

After the transition rates have been calculated by (2) to (5), the balance equations in (6) and (7) can be set up for each scheme. By solving the linear equations, the steady-state probability for each state of each scheme can be calculated. From the steady-state probabilities, transition rates, and unit costs, the total cost for each scheme can be calculated by (8). Based on these results, the trends of costs against the variation in parameters and unit costs can be analyzed. Finally, the performance of different schemes can be evaluated by comparing the total costs.

4.2 Analysis

The performance of the four LS schemes in terms of individual cost groups and the total cost (C_T) has been analyzed by using the model. The cost groups include the penalty cost (C_P), conversation cost (C_C), location management cost (C_L), and switching cost (C_S). By taking the levels as independent systems and the switching of MT between levels not being considered, the performance between individual levels is also analyzed. The above two types of analysis are performed against the variation in the values of the parameters in (1), the unit costs and cost groups in Fig. 4,

TABLE 3
Results for Basic Set of Parameter Values:
(a) Cost Group Values, (b) Steady-State Probability Distributions, (c) Basic Set of Parameter Values

Cost (cost unit per second)		NLS	ALS	RLS	SLS	Level 1	Level 2	Level 3	Level 4
penalty	C_P	10.090	2.502	2.502	2.502	10.090	350.045	713.610	995.294
conversation	C_C	19.844	6.507	19.801	6.810	19.844	2.644	0.254	0.001
location management	C_L	10.081	3.299	10.058	4.692	10.081	1.366	0.157	0.002
switching	C_S	0.000	0.594	0.001	0.199	0.000	0.000	0.000	0.000
total	C_T	40.015	12.902	32.362	14.203	40.015	354.055	714.021	995.297

(a)

Registered level (status)	state	NLS	ALS	RLS	SLS	Level 1	Level 2	Level 3	Level 4
0	π_{00}	0.0100	0.0023	0.0023	0.0023	0.0100	0.3400	0.6700	0.9900
1 (idle)	π_{01}	0.9000	0.2030	0.8961	0.3361	0.9000			
1 (conversation)	π_{11}	0.0900	0.0203	0.0896	0.0225	0.0900			
2 (idle)	π_{02}	0.0000	0.3980	0.0095	0.4509		0.6009		
2 (conversation)	π_{12}	0.0000	0.0398	0.0009	0.0421		0.0591		
3 (idle)	π_{03}	0.0000	0.2970	0.0015	0.1192			0.3040	
3 (conversation)	π_{13}	0.0000	0.0297	0.0001	0.0256			0.0261	
4 (idle)	π_{04}	0.0000	0.0090	0.0000	0.0009				0.0096
4 (conversation)	π_{14}	0.0000	0.0009	0.0000	0.0004				0.0004

(b)

Variable type	Parameter	Level 1	Level 2	Level 3	Level 4
structure variable	L	4			
time variable	δ (second)	1000			
	τ (second)	100			
	σ (second)	10000	2000	400	80
probability variable	β	0.99	0.66	0.33	0.01
cost variable	c_{cv} (cost unit)	100	20	4	0.8
	c_{lm} (cost unit)	10	2	0.4	0.08
	c_{ch} (cost unit)	5000	1000	200	40
	c_{cs} (cost unit)	10000	2000	400	80
	c_{ct} (cost unit)	2000	400	80	16
	c_{ls} (cost unit)	1000	200	40	8
	c_{lt} (cost unit)	200	40	8	1.6
	c_{sct} (cost unit)	4000	800	160	32
	c_{sco} (cost unit)	4000	800	160	32
	c_{sit} (cost unit)	400	80	16	3.2
	c_{sio} (cost unit)	400	80	16	3.2
	c_r (cost unit)	1000000			
	c_d (cost unit)	1000000			

(c)

and the ratios between values of level-dependent parameters (e.g., $\frac{\sigma_i}{\sigma_{i+1}}$).

We have used many different sets of parameters and unit costs for the analysis. We have found that, although the absolute performance between the levels or schemes depends on the set of values being used, the trends of the performance are not affected by the selection of input values. For example, the trend of C_T for the ALS scheme against the variation in δ is independent of the values of other parameters such as τ. A system is then defined for analyzing the performance trends of individual levels and the performance comparison of various schemes under different conditions.

A system with four levels is used in the analysis. For example, Level 1 may be the mobile satellite, Level 2 may be the cellular with macrocell, Level 3 may be the cellular with microcell, and Level 4 may be the cordless system. The

results of the costs and steady-state probability distributions obtained using a basic set of parameter values and unit cost values are shown in Table 3. The following are some considerations in selecting the input values for constructing the model: β_i is set to near unity for the first level, which represents the mobile satellite system with global coverage, and decreases for subsequent lower levels. The unit call-rejection penalty cost (c_r) and the unit call-dropout penalty cost (c_d) are set to be very large as compared to the level-dependent costs to reflect the importance of being able to receive or continue a call. In assigning values to the time variables and cost variables, their absolute values are not so important when compared with their relative values. In contrast, the absolute values of the probability variables, which take the values between 0 and 1, are important in addition to their relative values among levels.

TABLE 4
Summary of the Performance Analysis Results

Increase in variable	Trend of total cost	Lowest cost scheme (with level switching)	Lowest cost level (level switching not considered)
mean time between call	↘	ALS	1
average conversation time per call	↗	ALS	1
cell residence time	↘	SLS → ALS	1
penalty cost	↗ NLS / others	NLS → ALS	4 → 3 → 2 → 1
conversation cost	↗ others / NLS	ALS → NLS	1 → 2 → 3 → 4
location management cost	↗ others / NLS	ALS → NLS	1 → 2 → 3 → 4
switching cost	↗ others / NLS	ALS → SLS → RLS → NLS	1
ratio between cell residence time in each level	↗	ALS → SLS	1
ratio between probability of under coverage in each level	↗	ALS	4 → 3 → 2 → 1
ratio between level-dependent costs in each level	↗	RLS → ALS → NLS	1 → 2 → 3 → 4

Table 4 summarizes the results of the performance analysis. It shows the trends of C_T of various LS schemes against the increases in the 10 variables. It also shows which scheme has the best performance under different conditions. In addition, it shows which level has the lowest cost in different situations when the switching of MT between levels is not considered. This level is the one selected by the NLS scheme for registration. The analysis shows that the relative performance in terms of the ratio of C_T among the schemes (i.e., $C_{T(NLS)} : C_{T(ALS)} : C_{T(RLS)} : C_{T(SLS)}$) or levels (i.e., $C_{T(\text{level } 1)} : C_{T(\text{level } 2)} : C_{T(\text{level } 3)} : C_{T(\text{level } 4)}$) is quite independent of the time variables such as δ, τ, and σ_i. The relative performance mainly depends on the cost variables like the unit penalty cost (c_p), unit conversation cost (c_c), unit location management cost (c_l), and unit switching cost (c_s). It is also found that the system behaves similarly for varying any one of the CV or LM cost components. The system also exhibits similar behavior when any one of the switching cost components is varied.

In studying the performance of individual levels against the increase in c_p, the level with the lowest C_T changes from Level 4 to Level 3, then to Level 2, and finally to Level 1 as shown in Table 4. As $\beta_4 < \beta_1$, the probabilities of call rejection and dropout are higher in Level 4 so that Level 4 is more sensitive to c_p than Level 1. When the value of c_p is small, Level 4 has the lowest C_T since the unit level-dependent costs are the lowest. When the value of c_p is large, Level 1 has the lowest C_T since β_1 is the highest. So, in the NLS scheme without any LS, the MT will select Level 4 when c_p is low and Level 1 when c_p is high. It results in four line segments with decreasing gradient when C_T of the NLS scheme is plotted against c_p, as shown in Fig. 9. Each line segment represents a different level selected by the MT. From the figure, C_T of the other schemes increase linearly with c_p. The NLS scheme is more sensitive to c_p than other schemes with LS. It is because the call rejection and dropout penalties in the NLS scheme (in which only one level can be used) are much higher than those of the other schemes (in which all the levels can be used), which results in higher C_P. LS can then be considered as an effective way to reduce C_P by minimizing the probabilities of rejecting and dropping calls.
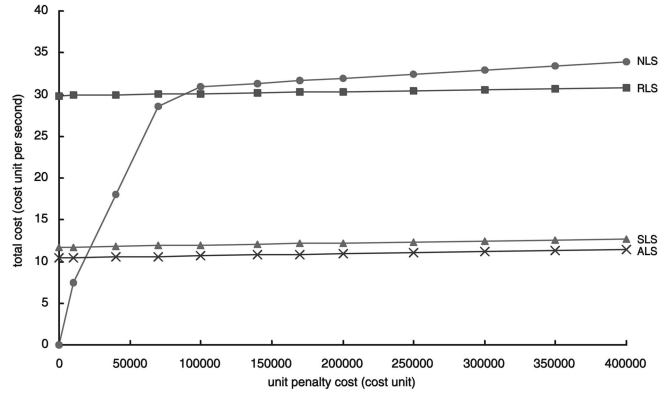

Fig. 9. Total costs of level switching schemes versus unit penalty cost.

It is also observed that, when the value of c_p is small, C_T of the NLS scheme is much less than those of the other schemes. It is because, when c_p is low, the MT will select Level 4 in the NLS scheme since the unit level-dependent costs are the lowest, although the probability of the MT being out of coverage is much higher due to lower β_4. As a result, the average time for LM and conversation is shorter than those of the other schemes as shown by the low values of steady-state probabilities of Level 4 for being idle or in conversation in Table 3b. It results in the low C_T of the NLS scheme due to the low C_C and C_L. However, it defeats the objective of the MCSs to provide communication services anywhere. So, we set c_p to very large values to show the importance of providing satisfactory services to the MUs.

In comparing the performance of individual levels against the increase in c_c and c_l, the level with the lowest C_T changes from Level 1 to Level 2, then to Level 3, and finally to Level 4, as shown in Table 4. Since c_c and c_l in Level 1 are much higher than those of the other levels, the increase in C_T for Level 1 will be greater than the other levels when c_c or c_l is increased and the level with the lowest c_c and c_l will have the best performance. The performance of the LS schemes in terms of C_T against c_c and c_l is shown in Fig. 10 and Fig. 11, respectively. C_T is found to increase linearly with c_c and c_l for all the schemes. The performance of the ALS and SLS schemes is found less sensitive to the variation of c_c and c_l than the NLS and RLS schemes. By examining the cost components of each scheme
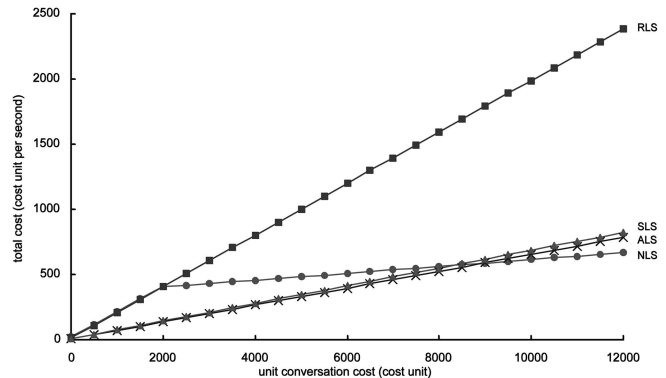

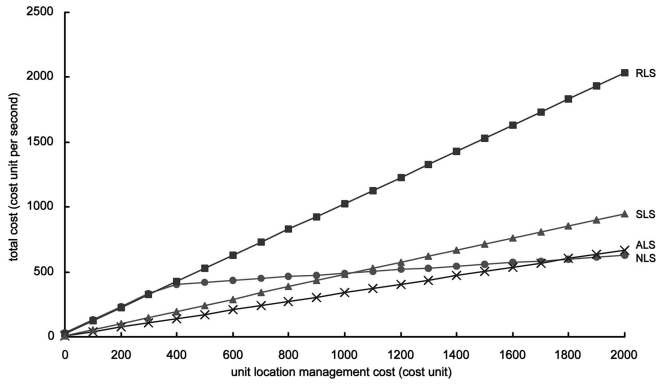Fig. 10. Total costs of level switching schemes versus unit conversation cost.

Fig. 11. Total costs of level switching schemes versus unit location management cost.

Fig. 13. State-transition diagram for a single-level system with channel availability considered.

in Table 3a, C_C and C_L are much lower in the ALS and SLS schemes than the NLS and RLS schemes. It proves that active LS can effectively reduce C_C and C_L by switching to a level with lower unit level-dependent costs at appropriate events. The performance of the ALS scheme is better than the SLS scheme in reducing C_C and C_L more effectively by performing LS more frequently.

The effect of c_s on the performance of the LS schemes has been studied and the results are shown in Fig. 12. Since there is no LS activity in the NLS scheme, C_T of the NLS scheme is not affected by c_s. For all the other schemes with LS, C_T is found to increase linearly with c_s. The increasing rate of C_T depends on the frequency of LS activities of the corresponding scheme. Since the ALS scheme has the highest LS activities, it is more sensitive to c_s than the other schemes. So, when c_s increases, the scheme with the lowest cost changes from ALS to SLS, then to RLS and, finally, to NLS, which is in the sequence of decreasing LS activities. It is noticed that C_S is insignificant with respect to other costs as shown in Table 3a. The exceptional case is when the value of c_s is extremely large. Thus, LS can be considered as an effective way to reduce the total COMM cost with very little cost in performing the LS activities.
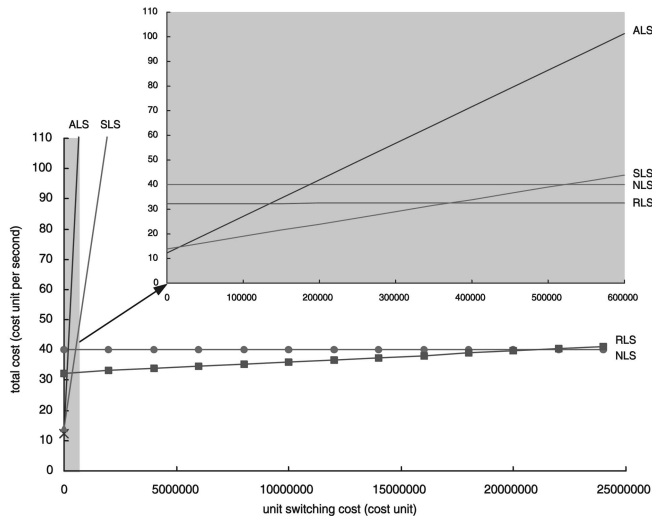

Fig. 12. Total costs of level switching schemes versus unit switching cost. (The shaded area is enlarged for easy reference.)
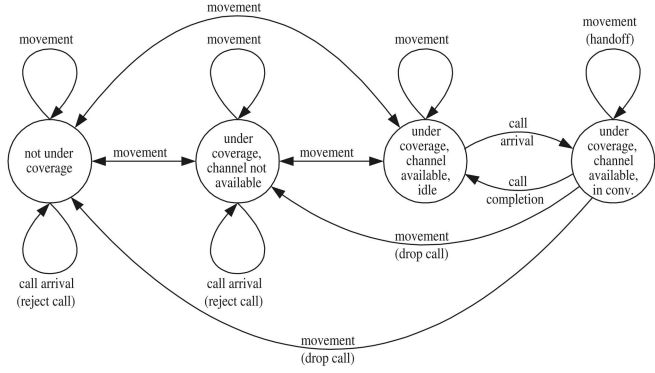
Finally, the performance of individual levels has been studied against the changes in the ratios between the level-dependent parameters of each level. First, when β_i of all the levels are equal, C_P for all the levels will be the same. The level with the lowest c_c and c_l (i.e., Level 4) will have the lowest C_T also. Second, when β_4 is much smaller than β_1, C_P will become more significant than C_C and C_L in Level 4. Then, Level 1 will have lower C_T due to higher β_1, although c_c and c_l are the highest. Similarly, when all the levels have the same unit level-dependent costs, Level 1 will have the lowest C_T due to lower C_P because of larger β_1. When the differences between the unit level-dependent costs of the levels are very large, Level 4 will have the lowest cost since the reduction in C_C and C_L will outweigh the effect of C_P.

5   ENHANCEMENT OF THE MODEL

In the previous sections, we assume the MTs can always find unoccupied channels to communicate with the MCSs. In this section, the effect of unavailability of channels for conversation will be investigated.

5.1   Performance Model

The performance model defined in Section 2 is enhanced to include one more main entity, namely, *channel*. The *probability of channel availability* is defined to be the probability that there are unoccupied channels available for the MTs to communicate with the MCSs. It is assumed to be constant for each level. When the MT moves among the cells in a certain level, it will be in one of the following three statuses instead of two: *not under coverage*, *under coverage but no channel available*, and *under coverage with available channel*. The status is determined by the *probability of under coverage* and the *probability of channel availability* of the particular level. For a single-level system, we now have four states instead of three, as shown in the state-transition diagram in Fig. 13.

The structure of the enhanced model is shown in Fig. 14. There are two types of cells, one with available channels and the other without available channels. It is assumed that LM can still be carried out in the cells without available channels. So, when the MT moves out of the coverage of Level 2 at I, it can switch to Level 3 at J instead of Level 1 for the system to continue the tracking of its location at a lower cost. However, channels must be available for starting or continuing a conversation. When
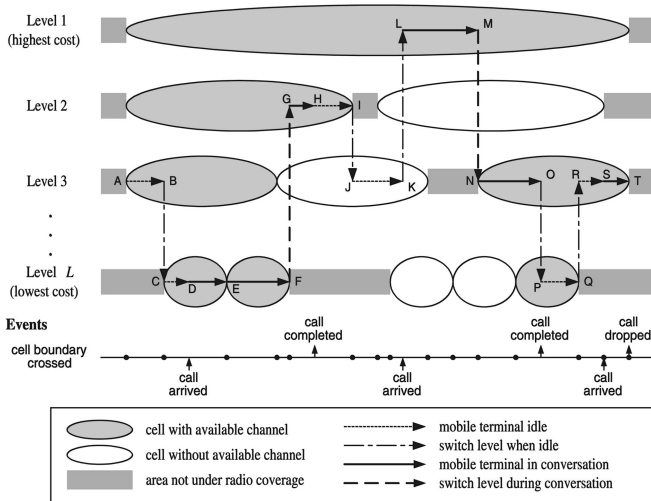
Fig. 14. Enhanced performance model structure.

the MT receives an incoming call when it moves to K, there is no channel available in the cell of Level 3 in which it resides. It has to switch to Level 1 to start the conversation at L since Level 1 is the only level with available channels, although the cost is the highest.

The number of states of the stochastic model, $S_{xyz_1 z_2 z_3 \cdots z_i \cdots z_L}$, increases since there will be one more status for z_i.

z_i = status in Level i $(i = 1 \ldots L)$

$$= \begin{cases} 0 & \text{not under coverage} \\ 1 & \text{under coverage, channel not available} \\ 2 & \text{under coverage, channel available,} \end{cases}$$

y = registered level

$$= \begin{cases} 0 & \text{if} & z_i = 0 \quad \forall \, i = 1 \ldots L \\ 1 \ldots L & \text{given} & \begin{cases} z_y \geq 1 & \text{when} & x = 0 \\ z_y = 2 & \text{when} & x = 1. \end{cases} \end{cases}$$

An additional parameter γ_i (the probability of channel availability in Level i) is added to (1) for the enhanced model. The rate of moving under coverage in Level i in (5) is split into the following two rates:

$\alpha_i \beta_i (1 - \gamma_i)$ = rate of moving under coverage in Level i (channel not available),

$\alpha_i \beta_i \gamma_i$ = rate of moving under coverage in Level i (channel available).

The state-transition diagram for a single-level system (Level i) in Fig. 13 can be transformed to the transition-rate diagram in Fig. 15.

In the cost model, call rejection and call dropout will also result if there is no channel available in the cells of all levels where the MT currently resides. The transitions and activities in Table 1 are modified to those in Table 5.

5.2 Level Switching Schemes

In the LS schemes, the determination of the target level is based on the resources required, which are now different for LM and for conversation. The resource required for LM is radio coverage while the resources required for conversation
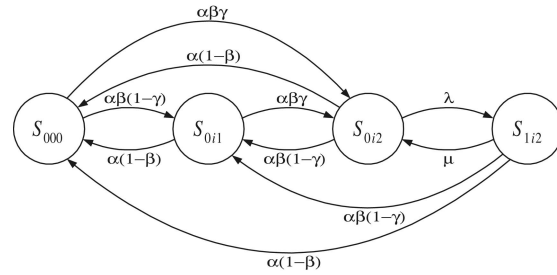
Fig. 15. Transition-rate diagram for single-level system with channel availability considered. (For Level i, $\alpha = \alpha_i$, $\beta = \beta_i$, and $\gamma = \gamma_i$.)

are both radio coverage and available channel in the cell where the MT currently resides. In the SLS scheme, LS can be based on movement-related events such as cell boundary crossing in addition to call-related events such as call arrival and completion. We have the *movement-based SLS scheme (MLS)* for the former case and the *call-based SLS scheme (CLS)* for the latter case, as shown in Table 6.

5.3 Performance Analysis

The performance of the five LS schemes in Table 6 has been analyzed using the enhanced model in the same way as in Section 4. The total number of states for each scheme can be calculated by (10), whose derivations can be found in [31].

$$\begin{aligned} n'_{\text{NLS}} &= 4L, \\ n'_{\text{ALS}} &= 2(3^L) - 2^L, \\ n'_{\text{RLS}} = n'_{\text{CLS}} &= 1 + L(3^L), \\ n'_{\text{MLS}} &= \frac{1}{2}[5(3^L) + 1] - 2^{L+1}. \end{aligned} \tag{10}$$

The effect of γ_i on the performance of the LS schemes has been studied. γ_i is set to unity for the lowest level, which means that the channel is always available in the private cordless system, and decreases for subsequent higher levels. $\gamma_1 = 0.7$, $\gamma_2 = 0.8$, $\gamma_3 = 0.9$, and $\gamma_4 = 1$ are added to the basic set of parameter values in Table 3. We have used different sets of γ_i for the performance analysis. We have found that the trends of the performance of the LS schemes against the variation of c_p, c_c, c_l, and c_s are quite independent of γ_i as shown in Figs. 16, 17, 18, and 19, respectively, which are similar to Figs. 9, 10, 11, and 12. The results show that channel availability has an insignificant effect on the trends of the performance of the LS schemes. Also, the chance of channel being unavailable for conversation is very rare in the current MCSs. The simple performance model defined in Section 2 is good enough for studying the performance trends of the LS schemes.

6 CONCLUSIONS

In this paper, a model is developed to represent the multitier environment of mobile communication systems as a single system with different levels of cells, each level representing a mobile communication system. A continuous-time stochastic process model parameterized by a set of state-transition rates and costs is constructed. The state-transition rates are derived from characteristics of

TABLE 5
Types of Transitions and Activities with Channel Availability Considered

Transition type	$T_a \in A_i$	$T_b \in B_j$	r_{ab}	d_{ab}
setup call	$x = 0$ $y = i > 0$ $z_i > 0$	$x' = 1$ $y' = j > 0$ $z'_j = 2$	λ	$c_{lt(i)} + c_{cs(j)}$
complete call	$x = 1$ $y = i > 0$ $z_i = 2$	$x' = 0$ $y' = j > 0$ $z'_j > 0$	μ	$c_{ct(i)} + c_{ls(j)}$
switch level when idle	$x = 0$ $y = i > 0$ $z_i > 0$	$x' = 0$ $y' = j \neq i$ $z'_j > 0$	$z'_j = 1: \quad \alpha_j \beta_j (1 - \gamma_j)$ $z'_j = 2: \quad \alpha_j \beta_j \gamma_j$	$c_{sio(i)} + c_{sit(j)}$
switch level during conversation	$x = 1$ $y = i > 0$ $z_i = 2$	$x' = 1$ $y' = j \neq i$ $z'_j = 2$	$\alpha_j \beta_j \gamma_j$	$c_{sco(i)} + c_{sct(j)}$
move under coverage from out of coverage	$x = 0$ $y = 0$ $z_i = 0 \quad \forall i$	$x' = 0$ $y' = j > 0$ $z'_j > 0$	$z'_j = 1: \quad \alpha_j \beta_j (1 - \gamma_j)$ $z'_j = 2: \quad \alpha_j \beta_j \gamma_j$	$c_{ls(j)}$
move out of coverage when idle	$x = 0$ $y = i > 0$ $z_i > 0$	$x' = 0$ $y' = 0$ $z'_j = 0 \quad \forall j$	$\alpha_i (1 - \beta_i)$	0
move out of coverage during conversation	$x = 1$ $y = i > 0$ $z_i = 2$	$x' = 0$ $y' = 0$ $z'_j = 0 \quad \forall j$	$\alpha_i (1 - \beta_i)$	$c_{ct(i)} + c_d$
move into cell without available channel during conversation	$x = 1$ $y = i > 0$ $z_i = 2$	$x' = 0$ $y' = j > 0$ $z'_j = 1$	$\alpha_i \beta_i (1 - \gamma_i)$	$c_{ct(i)} + c_{ls(j)} + c_d$

Activity type	$T_a \in A_i$	r_a	d_a
conversation	$x = 1$ $y = i > 0$ $z_i = 2$	1	$c_{cv(i)}$
location management	$x = 0$ $y = i > 0$ $z_i > 0$	1	$c_{lm(i)}$
handoff	$x = 1$ $y = i > 0$ $z_i = 2$	$\alpha_i \beta_i \gamma_i$	$c_{ch(i)}$
call rejection	$x = 0$ $y = i$ $z_j < 2 \quad \forall j$	λ	c_r

mobile communication systems (e.g., coverage area) and behavior of mobile users (e.g., calling pattern and mobility). The costs include those of staying in certain states and moving between states. The probabilities of staying in certain states and rates of moving between states are derived by using steady-state analysis. These probabilities and rates are used to calculate the total cost of the system for performance evaluation.

Several level switching schemes based on different triggering events are proposed in this paper for optimizing the total cost in the multitier system. The performance of these schemes is analyzed and compared by using the model developed. From this analysis, we have obtained the behavior of the schemes with respect to variations of the characteristics of mobile systems and mobile users. Also from the comparison, we can find the switching scheme with the lowest cost under any given conditions.

The relative performance among various level switching schemes mainly depends on the unit costs such as the penalty, conversation, location management, and switching costs. The No Level Switching scheme only uses one mobile

system, so it is sensitive to the penalty costs due to high probabilities of rejecting and dropping calls. The Reactive Level Switching scheme uses level switching to reduce the penalty costs but not the conversation or location management costs, so it is sensitive to both the conversation and location management costs. The Aggressive Level Switch-

TABLE 6
Level Switching Schemes and Triggering Events with
Channel Availability Considered

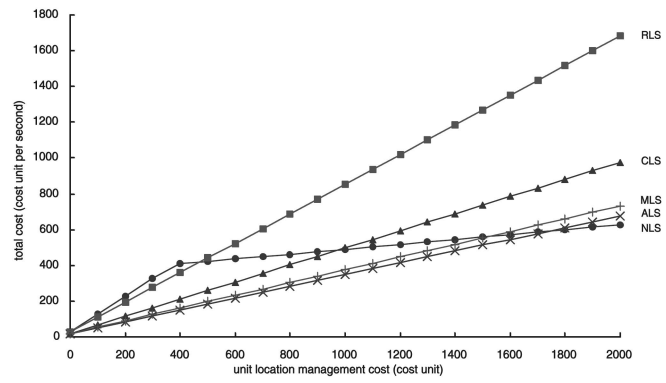| Perform level switching ? / Event / Level Switching Scheme		Call arrival and completion (call-related)		Cell boundary crossing (movement-related)	
		Resource available in current level	Resource not available in current level	Resource available in current level	Resource not available in current level
No Level Switching	NLS	No	No	No	No
Aggressive Level Switching	ALS	Yes	Yes	Yes	Yes
Reactive Level Switching	RLS	No	Yes	No	Yes
Call-based Selective Level Switching	CLS	Yes	Yes	No	Yes
Movement-based Selective Level Switching	MLS	No	Yes	Yes	Yes
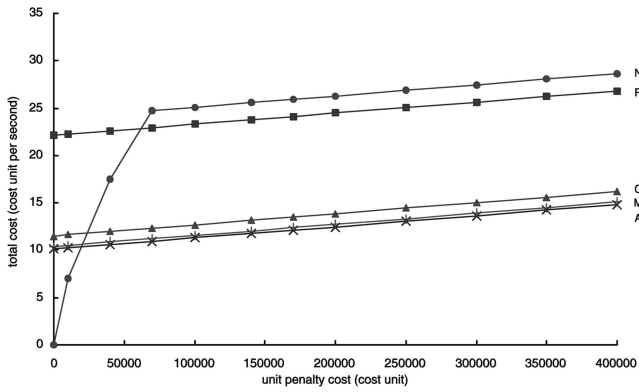
Fig. 16. Total costs of level switching schemes versus unit penalty cost (with channel availability considered).

ing scheme utilizes all possible level switching events to reduce the conversation and location management costs in addition to the penalty costs, so it is sensitive to the switching costs. The Selective Level Switching scheme balances the switching costs with the conversation and location management costs by utilizing part of the switching events. By using the model, the most cost-effective scheme can be selected to match the characteristics of individual mobile users and mobile systems for optimizing overall system performance.

Level switching is found to be an effective way to minimize call-rejection and call-dropout activities. It also effectively reduces the conversation and location management costs. Although additional costs are introduced by level switching, they are found to be insignificant in most cases, except when unit switching costs are extremely large. So, performance of the system can be improved in general by increasing level switching activities to reduce the conversation and location management costs.

The performance model has been enhanced to investigate the effect of channel availability on the level switching schemes. It is found that this factor has no significant impact on the trends of the performance of the switching schemes against the variation of the cost variables. It justifies that the basic performance model with channels always available is a good approximation for studying performance trends of the level switching schemes.

Fig. 18. Total costs of level switching schemes versus unit location management cost (with channel availability considered).

Although current technologies may not be mature enough to support real-time switching between mobile communication systems with different air-interface or networking standards in a multitier environment, the performance model can still be applied to switching activities between systems with the same standard but with different characteristics like cell sizes and radio coverage (e.g., GSM systems operating at 900 MHz, 1,800 MHz and 1,900 MHz). Furthermore, the performance model developed for mobile communication systems can be applied to other multitier systems by generalizing the model. It is done by defining the generalized terms for the characteristics of the model, which include the entities, events, status, parameters, and costs of the model [31]. Future works on the performance model include enhancements to provide for channel availability not being constant and to extend its applicability to more types of multitier systems. For example, the circuit-switching model adopted from the mobile communication system for voice communication can be extended to a packet-switching model for data communication in which requests for resources will be queued up instead of being rejected.
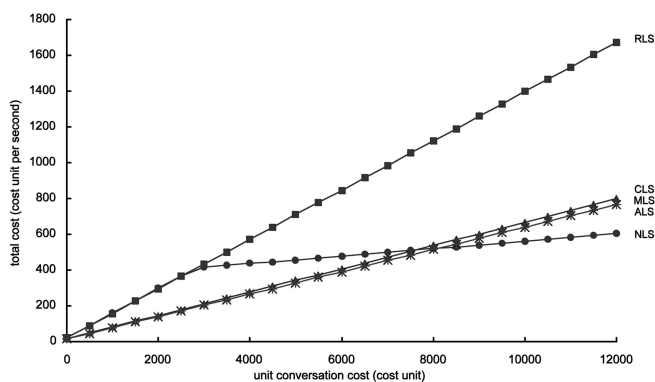
ACKNOWLEDGMENTS

Fig. 17. Total costs of level switching schemes versus unit conversation cost (with channel availability considered).
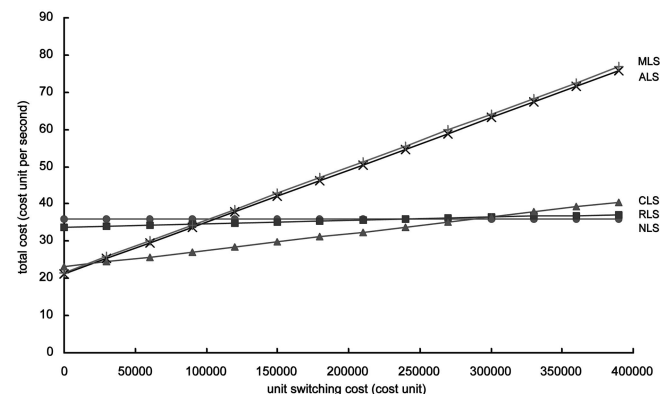

Fig. 19. Total costs of level switching schemes versus unit switching cost (with channel availability considered).

REFERENCES

[1] M. Zeng, A. Annamalai, and V.K. Bhargava, "Recent Advances in Cellular Wireless Communications," *IEEE Comm. Magazine,* vol. 37, no. 9, pp. 128-138, Sept. 1999.

[2] A. Ganz, C.M. Krishna, D. Tang, and Z.J. Haas, "On Optimal Design of Multitier Wireless Cellular Systems," *IEEE Comm. Magazine,* vol. 35, no. 2, pp. 88-93, Feb. 1997.

[3] X. Lagrange, "Multitier Cell Design," *IEEE Comm. Magazine,* vol. 35, no. 8, pp. 60-64, Aug. 1997.

[4] I. Corden and A. Kokkos, "Mobile Systems—The Next Generation," *Proc. IEE Colloquium Mobile Comm. towards the Next Millennium and Beyond,* pp. 5/1-5/9, 1996.

[5] I.F. Akyildiz, J. McNair, J. Ho, H. Uzunalioğlu, and W. Wang, "Mobility Management in Current and Future Communications Networks," *IEEE Network,* pp. 39-49, July 1998.

[6] K. Buchanan, R. Fudge, D. McFarlane, T. Philips, A. Sasaki, and H. Xia, "IMT-2000: Service Provider's Perspective," *IEEE Personal Comm.,* pp. 8-13, Aug. 1997.

[7] N. Faggion and T. Hua, "Personal Communications Services through the Evolution of Fixed and Mobile Communications and the Intelligent Network Concept," *IEEE Network,* pp. 11-18, July 1998.

[8] I.F. Akyildiz, J. McNair, J.S.M. Ho, H. Uzunalioğlu, and W. Wang, "Mobility Management in Next-Generation Wireless Systems," *Proc. IEEE,* vol. 87, no. 8, pp. 1345-1384, 1999.

[9] A.U.H. Sheikh, *Wireless Communications: Theory and Techniques.* Kluwer Academic, 2004.

[10] Y.-H. Wang, C.-H. Tsai, and C.-C. Chuang, "Construction of IP-based Multi-Tier Network for Mobile Multimedia Communication Services," *Proc. Int'l Conf. Parallel Processing,* pp. 465-470, 2001.

[11] A. Iera, A. Molinaro, and S. Marano, "Traffic and Resource Management in Multi-Tier Networks Exploiting Satellite Coverage," *Proc. Wireless Comm. and Networking Conf.,* vol. 1, pp. 169-173, Sept. 1999.

[12] A. Iera, A. Modafferi, and A. Molinaro, "Access Control and Handoff Management in Multi-Tier Multimedia Wireless Systems," *Proc. Wireless Comm. and Networking Conf.,* vol. 3, pp. 1518-1522, Sept. 1999.

[13] Y.-H. Wang, C.-H. Tsai, and H.-M. Huang, "Mobility Management of IP-Based Multi-Tier Network Supporting Mobile Multimedia Communication Services," *Proc. Int'l Conf. Distributed Computing Systems Workshops,* pp. 347-352, July 2002.

[14] S.G. Niri and R. Tafazolli, "Cordless-Cellular Network Integration for the 3rd Generation Personal Communication Systems," *Proc. IEEE Vehicular Technology Conf.,* vol. 1, pp. 402-408, 1998.

[15] W. Zhao, R. Tafazolli, and B.G. Evans, "An Integrated System Architecture for GSM and Satellite PCN," *Proc. IEE Colloquium Integration of Satellite and Terrestrial PCS,* pp. 3/1-3/9, 1997.

[16] L.-R. Hu and S. Rappaport, "Adaptive Location Management Scheme for Global Personal Communications," *Proc. IEE Comm.,* vol. 144, no. 1, pp. 54-60, 1997.

[17] Y. Chung, D.-J. Lee, B.-C. Shin, and D.-H. Cho, "A New Macrocell/Microcell Selection Method in Multitier Cellular System," *Proc. IEEE Vehicular Technology Conf.,* vol. 2, pp. 1016-1020, 1999.

[18] Y. Chung, D.-J. Lee, D.-H. Cho, and B.-C. Shin, "Macrocell/Microcell Selection Schemes Based on a New Velocity Estimation in Multitier Cellular System," *IEEE Trans. Vehicular Technology,* vol. 51, no. 5, pp. 893-903, Sept. 2002.

[19] A. Iera, A. Molinaro, and S. Marano, "Wireless Broadband Applications: The Teleservice Model and Adaptive QoS Provisioning," *IEEE Comm. Magazine,* pp. 71-75, Oct. 1999.

[20] A. Iera, A. Molinaro, and S. Marano, "Managing Symmetric and Asymmetric Data Traffic in Integrated Terrestrial-Cellular and Satellite Systems," *IEEE Personal Comm.,* vol. 7, no. 2, pp. 56-64, Apr. 2000.

[21] Y. Fang, "Registration Traffic and Service Availability for Two-Tier Wireless Networks," *Proc. Wireless Comm. and Networking Conf.,* vol. 3, pp. 1090-1095, Sept. 2000.

[22] D. Gu and S.S. Rappaport, "A Dynamic Location Tracking Strategy for Mobile Communication Systems," *Proc. IEEE Vehicular Technology Conf.,* vol. 1, pp. 259-263, 1998.

[23] B. Suh, J.-S. Choi, and J.K. Kim, "Mobile Location Management Strategy with Implicit Location Registration," *Proc. IEEE Vehicular Technology Conf.,* vol. 3, pp. 2129-2133, 1999.

[24] Z. Naor and H. Levy, "Cell Identification Codes for Tracking Mobile Users," *Proc. IEEE INFOCOM* vol. 1, pp. 28-35, 1999.

[25] I.F. Akyildiz, J.S. M. Ho, and Y.-B. Lin, "Movement-Based Location Update and Selective Paging for PCS Networks," *IEEE/ACM Trans. Networking,* vol. 4, no. 4, pp. 629-638, 1996.

[26] L. Tassiulas and F.M. Anjum, "A Hierarchical Multiresolution Registration Structure for Mobility Tracking," *Proc. Int'l Conf. Universal Personal Comm.,* pp. 617-622, 1996.

[27] J.-L. Chen and J.-W. Wu, "Adaptive Location Management in Wireless Mobile Communications," *IEEE Int'l Conf. Comm.,* vol. 1 pp. 619-624, 1999.

[28] F.T. P. Camarda, G. Schiraldi, and R. Valla, "Mobility and Performance Modeling in Cellular Communication Networks," *Mobile Computing and Comm. Rev.,* vol. 1, no. 4, pp. 25-32, Oct. 1997.

[29] R.B. Cooper, *Introduction to Queueing Theory,* second ed. North Holland, 1981.

[30] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis,* third ed. McGraw Hill, 2000.

[31] C.K. Ng, "Location Management and Level Switching Schemes in Multitier Mobile Communication Systems," PhD thesis, Univ. of Hong Kong, 2005.
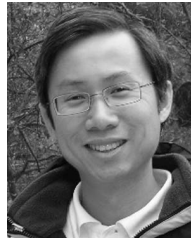
C.K. Ng received the BSc degree in electrical and electronic engineering and the MSc degree in computer science, and the PhD degree from The University of Hong Kong in 1987, 1993, and 2005, respectively. His research interests include location management and performance modeling of mobile communication systems.

H.W. Chan's earlier research works were in developing algorithms for combined routing and flow control in computer networks. In recent years, his research interests include network security, mobile computing, and their performance evaluation.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.