

An Efficient Handoff Management Scheme for Mobile Wireless ATM Networks

King-Sun Chan, Sammy Chan, King-Tim Ko, Kwan L. Yeung, *Senior Member, IEEE*, and Eric W. M. Wong

Abstract—In this paper, a new handoff management scheme for wireless ATM networks is proposed. In this scheme, all cells are connected to their neighboring cells by permanent virtual circuits (PVC's) and to the access switch (AS) by switched virtual circuits (SVC's) which are only for new calls. Some carefully chosen cells, called *rerouting cells*, are also connected to the AS by PVC's. In summary, if a mobile roams to an ordinary neighboring cell, its traffic path is simply elongated by a PVC connecting the old and new cells. If a mobile roams to a rerouting cell, its traffic path is rerouted to a PVC between the AS and rerouting cell. By using PVC's for handoff calls, we can guarantee fast and seamless handoff. At the same time, our scheme improves the path efficiency by limiting the maximum number of hops that a path can be extended. Also, allowing path rerouting at a suitable time means the network resources are more efficiently utilized.

Index Terms—Handoff, rerouting, wireless ATM.

I. INTRODUCTION

IN WIRED networks, ATM-based B-ISDN will be the single infrastructure which integrates all types of services in the same network. Recently, many research efforts have been devoted to extend the ATM capability into wireless networks, leading to the emergence of wireless ATM networks [1]. To support mobility in wireless ATM networks, fast and seamless handoff is a crucial issue. Due to the very high transmission speed, a short connection interruption will cause a large amount of information loss. As the population of mobiles increases, the cell radii will be further reduced. So the handoff would occur more frequently in the future. The handoff processing will be the bottleneck if centralized handoff processors are adopted.

According to [2], the handoff schemes proposed in the literature for wireless ATM networks can be divided into four categories: 1) full establishment; 2) connection extension; 3) incremental reestablishment; and 4) multicast establishment. In the full establishment scheme [3], each handoff triggers an establishment of a new virtual circuit (VC) between the terminals. This is rather time consuming and hence cannot guarantee seamless handoff required by wireless ATM networks. In the connection extension scheme [4], [5], for each handoff, the traffic path is elongated from the current cell to the new cell by appending one hop to the existing connection. This hop is a separate virtual connection between the current and new cells through the

wired network. The advantages of this scheme are that it is very simple, the cell sequence can be easily preserved, and the handoff process is very fast. However, it has a drawback that when a mobile makes a lot of handoffs, the path would be continuously elongated, which leads to path inefficiency. In the incremental reestablishment scheme [6], the traffic path would be partly reestablished in each handoff. Here, the main problem is still the time constraint. In the multicast establishment scheme [7], [8], whenever a new call is initiated, a set of virtual connections, named a *virtual connection tree* (VCT), is created, reaching all cells to which the mobile might move in the future. When the mobile roams to one of these cells, the preestablished VC can be used immediately. The scheme enables fast handoff and guarantees quality of service (QoS) even during handoff as network resources are preallocated. However, it consumes too much resources and results in network inefficiency.

In this paper, we propose a new handoff management scheme. In this scheme, some cells are defined as rerouting cells and others are ordinary cells. Each cell is connected to its neighboring cells by permanent virtual circuits (PVC's). An ordinary cell is connected to its local access switch (AS) only by switched virtual circuits (SVC's). A rerouting cell is connected to its AS by both SVC's and PVC's. PVC's are only used for handoff calls. In the linear environment, rerouting cells are regularly spaced. If a mobile roams to a rerouting cell, the traffic path would be rerouted to a PVC between the cell and AS. If a mobile roams to an ordinary cell, the traffic path is simply elongated by a PVC between the current and new cells. In the planar environment, seven cells are grouped as a cluster. The central cell is designated as a rerouting cell. A system parameter r is defined in this case. There is a handoff counter in each mobile. Each time the mobile makes a handoff, the counter is increased by one. Each time the traffic path is rerouted, the counter is reset to zero. If a mobile roams to a rerouting cell, its traffic path is rerouted to the rerouting cell immediately, as in the linear case. If a mobile roams to an ordinary cell and the handoff counter is less than the system parameter r , the traffic path is simply elongated by a PVC between the current and new cells. If the counter is equal to r , the traffic path would be rerouted to the new cell via the rerouting cell in the same cluster.

In essence, our scheme has the advantages of both the connection extension and incremental reestablishment schemes. On one hand, it enables fast handoff by path elongation using PVC, but at the same time, it improves the path efficiency by limiting the maximum number of hops that a path can be extended. On the other hand, it allows path rerouting to efficiently utilize the network resources. Also, since rerouting is restricted to certain types of handoff calls (handoff counter = r or rerouting cell

Manuscript received March 5, 1998; revised April 27, 1999. This work was supported by the Research Grants Council of Hong Kong (RGC) under Grant 9040205.

The authors are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong.

Publisher Item Identifier S 0018-9545(00)03690-2.

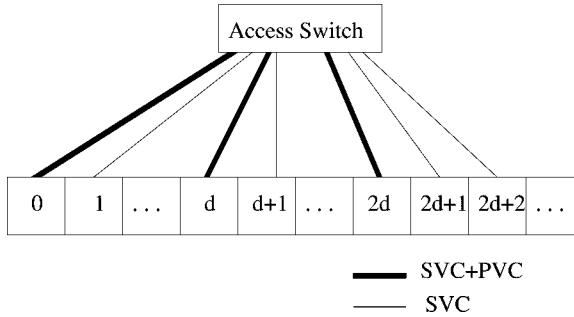


Fig. 1. The architecture of wireless ATM networks in linear environment.

is hit), the processing load is much less than that in the incremental reestablishment scheme. The remainder of this paper is organized as follows. In Section II, the proposed scheme in the linear environment is described in detail and its performance is evaluated. In Section III, we extend our scheme into the planar environment and analyze the performance. In Section IV, we explain our scheme in the general environment with multiple AS's. Finally, we conclude this paper in Section V.

II. LINEAR ENVIRONMENT

A. Scheme Description

Here, we consider the linear environment, as shown in Fig. 1. Some cells, say cells $0, d, 2d, 3d, \dots$, are assigned to be rerouting cells. Between the rerouting cells and AS, there are SVC's and PVC's. Between those ordinary cells and the AS, only SVC's can be established for new calls. Between a pair of neighboring cells, there are PVC's for handoff calls. If a mobile roams into a rerouting cell, the traffic path would be rerouted directly to the rerouting cell. If a mobile roams into an ordinary cell, the traffic path is elongated by a PVC between the current and new cells. Here, we gather all reroutings of handoff calls to some rerouting cells, thus, we obtain some statistical multiplexing gain and reduce the required number of PVC's between cells and the AS.

Now, we illustrate how the scheme operates by considering an example. Referring to Fig. 1, let the base station in cell i be denoted as BS_i . Assume d is four. When a mobile initiates a new call in cell 1, BS_1 will establish a SVC between itself and the called terminal through the AS. Then, a VCI is assigned to this call. Consider the mobile roams to its neighboring cell, say cell 2, before the call is finished. As shown by the signaling procedure depicted in Fig. 2, the mobile first sends a "handoff request" message to BS_2 , including the ID of BS_1 , the call's VCI, QoS requirements. Since cell 2 is an ordinary cell, BS_2 checks the availability of PVC's between itself and BS_1 to see whether it can support the call. If yes, it sends BS_1 a "PVC allocation" message including the identifier of the PVC to be used for path elongation. When BS_2 receives the acknowledgment for the PVC allocation from BS_1 , it sends a "handoff ready" message to the mobile. Then the mobile performs the wireless channel handoff, and after that, sends a "handoff" to BS_2 indicating the wireless channel handoff is completed. Upon receiving the "handoff" from the mobile, BS_2 sends the "handoff indication" to BS_1 . Immediately after receiving the "handoff

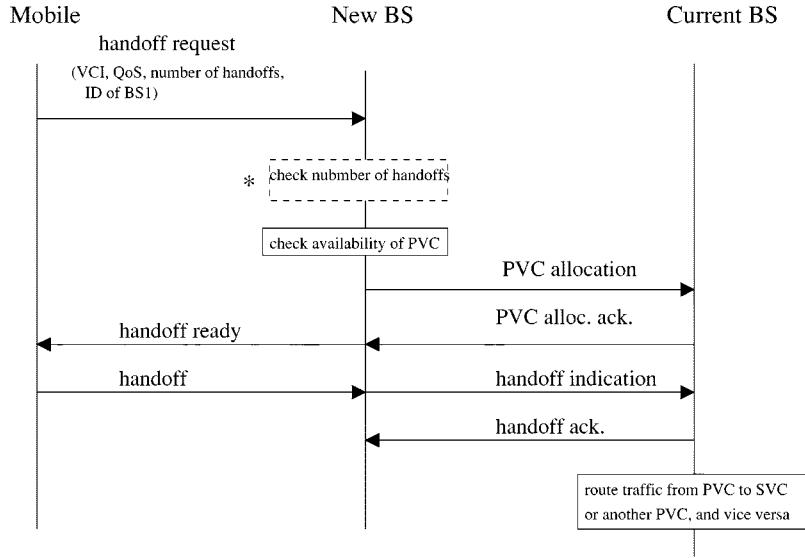
indication" message from BS_2 and returning the acknowledgement for the indication, BS_1 forwards the traffic coming from the SVC to the PVC and vice versa.

Then, consider the mobile goes on to visit cell 3. The signaling procedure is the same as that of the previous handoff. The traffic path between the AS and the mobile now becomes: AS- BS_1 - BS_2 - BS_3 -mobile, where between the AS and BS_1 is a SVC; between BS_1 and BS_2 , as well as BS_2 to BS_3 are PVC's.

Next, consider the mobile goes on to visit cell 4. Note that cell 4 is a rerouting cell, which means that the traffic path should be rerouted. The handoff signaling procedure is shown in Fig. 3. The mobile first sends a "handoff request" to BS_4 . BS_4 now checks if any PVC connecting to the AS is available. If that is true, BS_4 will send a "PVC allocation" message to the AS including the call's VCI and the identifier of the PVC to be used for the path rerouting. After receiving the "PVC allocation" from BS_4 , the AS sends an acknowledgment to that allocation. After receiving the ACK from the AS, BS_4 sends the "handoff ready" to the mobile. Then the mobile performs the wireless channel handoff, and after that, send a "handoff" to BS_4 indicating the wireless channel handoff is completed. Meanwhile, the mobile also sends a "PVC release" signal to BS_3 , informing it to release the reserved PVC. When BS_4 has received the "handoff" signal from mobile, it sends "handoff indication" to the AS. Then the AS will update the routing table and acknowledge the indication. At the same time, when BS_3 has received the "PVC release" signal from the mobile, it releases the PVC between itself and BS_2 and sends a similar message to BS_2 . BS_2 releases the reserved PVC and propagates the message to BS_1 . Then BS_1 will release the SVC connecting to the AS for the call.

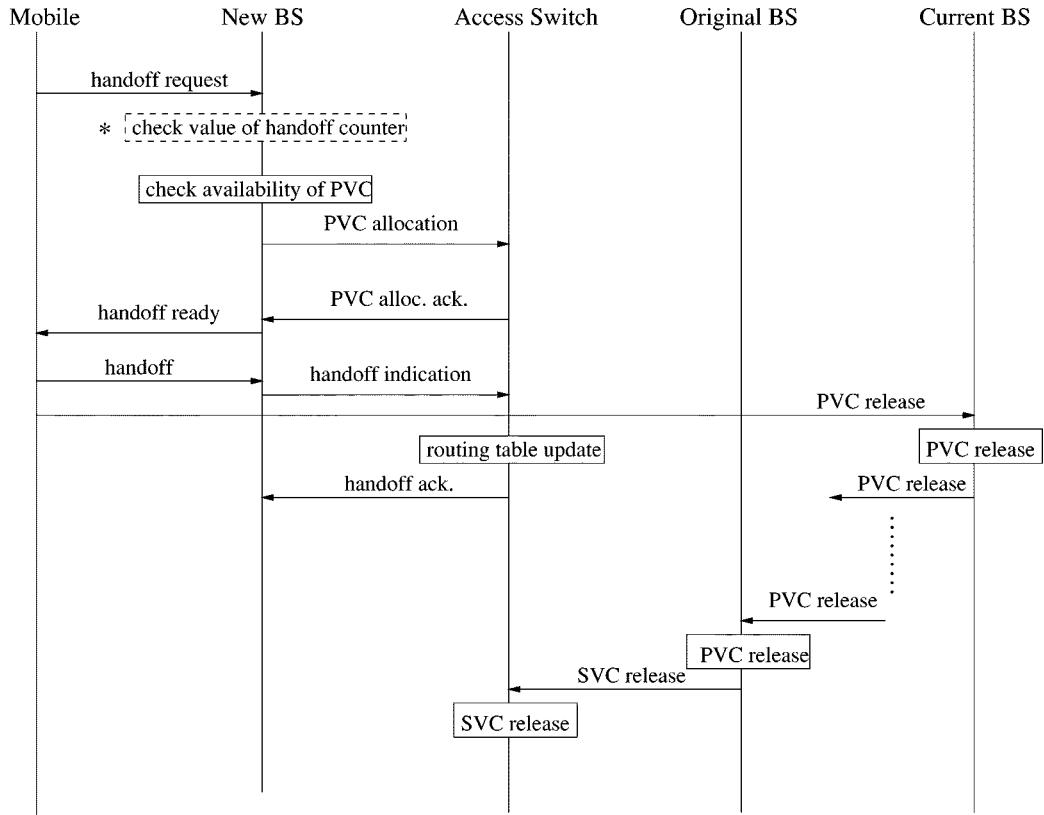
Since ATM is an connection-oriented technology, it is important to preserve the cell sequence during handoff. In our scheme, when the path is elongated after a handoff, the cell sequence is naturally preserved. Only when the path is rerouted, a mechanism is needed to ensure the integrity of cell sequence. Here, we provide a simple mechanism for this purpose. Referring to Fig. 3, we assume there is a buffer in the AS for each PVC connected to the rerouting cell. After the switch allocates a PVC to the rerouting cell upon the request of rerouting, it sends a special cell, called *downlink end-connection* (DEC) cell, to the mobile along the existing connection. At the same time, it stops transmitting subsequent cells to the mobile and keeps them in the buffer. Only when the mobile receives this DEC cell, then it can acknowledge the handoff indication from the rerouting cell. Once this acknowledgment reaches the AS, the switch will first release the buffered cells, then relay subsequent cells, to the mobile via the new path. By this mechanism, the cell sequence can easily be preserved. Similarly, to preserve the sequence of uplink traffic, as soon as the mobile initiates a handoff, it sends a *uplink end-connection* (UEC) cell along the old path to the AS. Subsequent uplink traffic are buffered in the mobile until the handoff is ready and the mobile has received the acknowledgement of the UEC cell from the AS.

To control the cell delay variation (CDV), the scheduling algorithm for jitter compensation [9] can be employed at the AS for downward traffic and base station for upward traffic, respectively. In this algorithm, the local delay time (LDT) is assigned



* This step is required only in the planar environment, as explained in Section 3.

Fig. 2. The signaling procedure for a handoff call using PVC connecting neighboring cells.



* This step is required only in the planar environment, as explained in Section 3.

Fig. 3. The signaling procedure of a handoff call applying for a PVC between the base station and AS.

for each node along the traffic path during the call setup phase. Cells delayed longer due to traffic path rerouting at the AS or base station have higher priority over normal cells and will be transmitted first. The cell delay variation can be improved by this algorithm.

To prevent looping in traffic path, the mobile needs to perform the loop detection. Each mobile is assumed to keep a path record to store the path information after the AS. For example, referring to Fig. 1, assume $d = 4$ and a mobile initiates a call in cell 1. The initial traffic path is: AS- BS_1 -mobile. The mobile stores

BS_1 in his path record. Then when the mobile moves to cell 2, BS_2 is added to the path record. Similarly when the mobile further moves to cell 3, BS_3 is added to the path record. Now in the path record, BS_1 , BS_2 , and BS_3 are stored, meaning the traffic path from AS to the mobile is $AS-BS_1-BS_2-BS_3$ -mobile. If the mobile now moves back to cell 2, since BS_2 is already in the path record, it knows that if another PVC between BS_3 and BS_2 is added to the path, there will be a loop on the traffic path: $BS_2-BS_3-BS_2$. So the mobile would request BS_2 to release the existing PVC between BS_2 and BS_3 and directly switch all traffic to it. After the PVC between BS_2 and BS_3 is removed, the mobile removes BS_3 from his path record, reflecting now the traffic path is: $AS-BS_1-BS_2$ -mobile. Considering another scenario in the above example, assume that after the mobile enters cell 3, it goes on to cell 4. As cell 4 is a rerouting cell, the traffic path will be rerouted. After the path rerouting, the path record of the mobile is updated to only contain BS_4 , reflecting now the traffic path is: $AS-BS_4$ -mobile. Then if the mobile goes back to cell 3, as the traffic path has been rerouted, no loop will be included along the path after a PVC between BS_4 and BS_3 is assigned to it, so the path can be extended by a PVC between BS_3 and BS_4 , and BS_3 is included in this path record.

B. Performance Analysis

Now we evaluate the performance of our scheme in linear environment. Here, we consider voice calls only. The following assumptions are given first.

- 1) The call holding time T_M is exponentially distributed with mean $1/\mu_M$.
- 2) The originating calls arriving in a cell follow a Poisson process with rate λ_0 .
- 3) The time interval R during which a mobile resides in a cell, called the cell sojourn time, has a general distribution, and the cell sojourn times, $R^{(1)}, R^{(2)}, \dots$, consecutively induced by movement of a mobile are independent and identically distributed.

When a mobile is within a cell, assume a VC labeled VC_1 connecting the cell's base station to the AS or to the base station of another cell is occupied for a ongoing call. The mobile will subsequently visit other cells, cells i_1, i_2, \dots . Assume cell i_k is a rerouting cell and cells i_j ($j < k$) are ordinary cells. When the mobile enters cell i_k , the traffic path of the call is rerouted if the call is still ongoing. The traffic path will only be elongated when the mobile enters cell i_j ($j < k$). We call VC_1 's holding time as T_k . The VC can be released when: 1) the mobile enters cell i_k and the traffic path is rerouted; 2) the connection is naturally terminated before the mobile enters cell i_k ; or 3) the connection is forced to be terminated due to handoff blocking when it moves to one of the cells i_1, i_2, \dots, i_k . No matter whether VC_1 is released due to cases 1)–3), we label the holding time of VC_1 as T_k . For example, referring to Fig. 1, assume $d = 4$. Consider that a mobile initiates a call in cell 1 when he is moving right and subsequently visits cells 2–5. When the mobile enters cell 2, a PVC labeled PVC_1 connecting the BS_2 to BS_1 is assigned for this call. When the mobile enters cell 4, as cell 4 is a rerouting cell, the call will be rerouted if the call is still ongoing.

So PVC_1 will be released in three cases: 1) if the call is naturally terminated before the mobile enters cell 4; 2) if the call is forced to be terminated when the mobile enters cell 3 or cell 4 due to handoff failure; and 3) when the mobile enters cell 4, the call is rerouted. We label the holding time of PVC_1 as T_2 , no matter whether PVC_1 is released due to cases 1)–3). Similarly, if a PVC labeled PVC_2 connecting BS_3 to BS_2 is assigned to this call, the holding time for PVC_2 is labeled as T_1 .

We now derive the mean of this random variable. First, we only consider cases 2 and 3. That is, the call will not be naturally terminated before the mobile makes the k th handoff attempt. Let p_f be the probability that the call would be blocked due to unavailability of PVC when the mobile tries to handoff to another cell. Let θ_k be the VC's holding time under this consideration, and n be the number of cells the mobile has resided during the interval in which the VC is occupied. For $k = 1$, we simply have

$$p(n = 1) = 1$$

and for $k > 1$, we have

$$\begin{aligned} P(n = 1) &= p_f \\ P(n = 2) &= (1 - p_f)p_f \\ &\dots \\ P(n = k - 1) &= (1 - p_f)^{k-2}p_f \\ P(n = k) &= (1 - p_f)^{k-1}. \end{aligned}$$

Let $N_k[z]$ be the generating function of variable n . We have

$$\begin{aligned} N_k[z] &= \sum_{i=1}^k z^i P(n = i) \\ &= \frac{p_f z + z^k (1 - p_f)^k (1 - z)}{1 - (1 - p_f)z}. \end{aligned}$$

Let $\Theta_k^*(s)$ be the Laplace–Stieltjes transform (LST) of $F_{\theta_k}(t)$, the distribution function of θ_k . We have

$$\Theta_k^*(s) = N_k[R^*(s)]$$

where $R^*(s)$ is the LST of the distribution function of R .

The distribution function of T_k is

$$F_{T_k}(t) = P(\min(\theta_k, T_M) \leq t).$$

Since call holding time is exponentially distributed

$$P(T_M \leq t) = 1 - e^{-\mu_M t}$$

so

$$F_{T_k}(t) = F_{\theta_k}(t) + (1 - F_{\theta_k}(t))(1 - e^{-\mu_M t}).$$

To obtain the mean of T_k , we first derive $T_k^*(s)$, the LST of $F_{T_k}(t)$. Let $F_{T_M}(t)$ be the distribution function of T_M and $f_{T_M}(t)$ be the density function. We have

$$\begin{aligned} F_{T_M}(t) &= 1 - e^{-\mu_M t} \\ f_{T_M}(t) &= \mu_M e^{-\mu_M t}. \end{aligned}$$

The LST of F_{T_M} is

$$T_M^*(s) = \int_0^\infty e^{-st} f_{T_M}(t) dt \\ = \frac{\mu_M}{s + \mu_M}.$$

Let $f_{T_k}(t)$ be the density function of T_k and $f_{\theta_k}(t)$ be the density function of θ_k . We have

$$f_{T_k}(t) = \frac{d}{dt} F_{T_k}(t) \\ = \frac{d}{dt} [F_{T_M}(t) + (1 - F_{T_M}(t))F_{\theta_k}(t)] \\ = \mu_M e^{-\mu_M t} + f_{\theta_k}(t) \\ - \mu_M e^{-\mu_M t} \int_0^t f_{\theta_k}(u) du + (1 - e^{-\mu_M t})f_{\theta_k}(t) \\ = \mu_M e^{-\mu_M t} + e^{-\mu_M t} f_{\theta_k}(t) - \mu_M e^{-\mu_M t} \int_0^t f_{\theta_k}(u) du.$$

So we have (1a), given at the bottom of the page.

With the assumption that R is exponentially distributed with mean μ_R [10], we have

$$R^*(s) = \frac{\mu_R}{s + \mu_R}$$

and hence

$$T_k^*(s) = \frac{\mu_M}{s + \mu_M} + \frac{s}{s + \mu_M} \\ \cdot \frac{\mu_R p_f + (s + \mu_M) \left[\frac{\mu_R(1 - p_f)}{s + \mu_M + \mu_R} \right]^k}{s + \mu_M + \mu_R p_f}.$$

Since

$$\frac{d}{ds} T_k^*(s) = -\frac{\mu_M}{(s + \mu_M)^2} \\ \cdot \left\{ 1 - \frac{\mu_R p_f + (s + \mu_M) \left[\frac{\mu_R(1 - p_f)}{s + \mu_M + \mu_R} \right]^k}{s + \mu_M + \mu_R p_f} \right\} \\ + A(s)$$

where $A(s)$ is a function of s with $A(0) = 0$ and the mean of T_k , $E[T_k]$, is given by

$$E[T_k] = -\frac{d}{ds} T_k^*(s)_{s=0} \\ = \frac{1 - \left[\frac{\mu_R(1 - p_f)}{\mu_M + \mu_R} \right]^k}{\mu_M + p_f \mu_R}.$$

As will be seen later, this result is used in dimensioning the VC quantities for our handoff scheme. Now, we will derive the expressions to be used for dimensioning. Here, we assume the mobile will only move along one direction. That is, if the mobile moves from left to right initially, then throughout the call the direction of movement will remain unchanged. Referring to Fig. 1, cells $0, d, 2d, \dots$ are designated as rerouting cells only in which can a handoff call be rerouted. Let n_{si} be the average of the number of occupied SVC's connecting the base station in cell i to AS for new calls ($0 \leq i \leq d-1$), n_{pi}^s be the number of required PVC's connecting the base station in cell i to AS for handoff calls (if cell i is not a special cell, then $n_{pi}^s = 0$), n_{pi}^l be the required number of PVC's connecting cell i to its left neighbor, and n_{pi}^r be the required PVC's connecting cell i to its right neighbor (if cell i is a special cell, $n_{pi}^l = n_{pi}^r = 0$).

To each cell, the handoff call arrival rate is given by [5]

$$\lambda_h = \frac{\mu_R(1 - p_0)\lambda_0}{\mu_M + \mu_R p_f}$$

where p_0 is the blocking probability of a new call.

If the probability that a handoff call is from the left neighboring cell is equal to that from the right neighboring cell, then the handoff call arrival rate from each side is $\lambda_h/2$.

To determine the required number of PVC's and SVC's in each cell, we only need to consider cell i , $0 \leq i < d$, because any cell j , $j \geq d$, has the same requirements as cell i if $j - \lfloor j/d \rfloor d = i$. We first derive the mean occupied SVC's in each cell. If a call is initiated and accepted in cell i , an SVC will be occupied by the call. Now consider $i \neq 0$. If the mobile moves to left, the mobile will reach cell 0 after going through i cells. The call will then be rerouted and the SVC in cell i is released. So the mean holding time is $E[T_i]$. On the other hand, if the user moves to right, the mean holding time is $E[T_{d-i}]$. Next, consider $i = 0$, the mobile will go through d cells before reaching a rerouting

$$T_k^*(s) = \int_0^\infty f_{T_k}(t) e^{-st} dt \\ = \frac{\mu_M}{s + \mu_M} + N_k[R^*(s + \mu_M)] - \frac{\mu_M}{s + \mu_M} N_k[R^*(s + \mu_M)] \\ = \frac{\mu_M}{s + \mu_M} + \frac{s}{s + \mu_M} N_k[R^*(s + \mu_M)] \\ = \frac{s}{s + \mu_M} \frac{R^*(s + \mu_M)p_f + (1 - p_f)^k R^*(s + \mu_M)^k (1 - R^*(s + \mu_M))}{1 - R^*(s + \mu_M)(1 - p_f)} + \frac{\mu_M}{s + \mu_M} \\ = \frac{\mu_M + [p_f s - \mu_M(1 - p_f)] R^*(s + \mu_M)}{(s + \mu_M)(1 - R^*(s + \mu_M)(1 - p_f))} + \frac{s(1 - p_f)^k (1 - R^*(s + \mu_M)) R^*(s + \mu_M)^k}{(s + \mu_M)(1 - R^*(s + \mu_M)(1 - p_f))} \quad (1a)$$

cell no matter which direction it moves, the mean holding time is thus $E[T_r]$. So we have

$$n_{si} = \lambda_0(1-p_0)(E[T_i] + E[T_{d-i}])/2, \quad \text{for } 0 < i < d \quad (1)$$

and

$$n_{s0} = \lambda_0(1-p_0)E[T_d]. \quad (2)$$

The average number of required SVC's per cell is

$$N_1 = \frac{\sum_{i=0}^{d-1} n_{si}}{d}.$$

Now let us derive the number of PVC's for cell i required to connect to its left neighboring cell (cell $i-1$) n_{pi}^l . If a mobile in cell i comes from $i-1$, then a PVC connecting cell i to cell $i-1$ will be occupied. In this case, the call arrival rate is $\lambda_h/2$ and the mean holding time is $E[T_{d-i}]$. So, we have

$$p_f = \frac{(\lambda_h E[T_{d-i}]/2)^{n_{pi}^l} / n_{pi}^l!}{\sum_{j=0}^{n_{pi}^l} (\lambda_h E[T_{d-i}]/2)^j / j!}. \quad (3)$$

On the other hand, if the mobile comes from right, the PVC's mean holding time is $E[T_i]$. So we have

$$p_f = \frac{(\lambda_h E[T_i]/2)^{n_{pi}^r} / n_{pi}^r!}{\sum_{j=0}^{n_{pi}^r} (\lambda_h E[T_i]/2)^j / j!}. \quad (4)$$

By (3) and (4), we can obtain n_{pi}^l and n_{pi}^r for a specified p_f . The average number of required PVC's connecting to neighboring cells per cell is

$$N_2 = \frac{\sum_{i=1}^{d-1} (n_{pi}^l + n_{pi}^r)}{d}.$$

For cell 0, only PVC's connected to the AS are required. Since the handoff call arrival rate is λ_h and the mean holding time is $E[T_d]$, we have

$$p_f = \frac{(\lambda_h E[T_d])^{n_{p0}^s} / n_{p0}^s!}{\sum_{j=0}^{n_{p0}^s} (\lambda_h E[T_d])^j / j!}. \quad (5)$$

The average number of required PVC's connecting to the AS per cell is

$$N_3 = \frac{n_{p0}^s}{d}.$$

Note that N_3 is calculated only for comparison with other schemes in the next section.

Define the processing load p_l to be the ratio of the number of handoffs whose traffic path need to be rerouted to the total number of handoffs. Then we have

$$p_l = \frac{1}{d}.$$

It can be seen that the processing load is inversely proportional to d .

C. Numerical Results

Now we present some numerical results, given the new call arrival rate is 11.9 calls per minute per cell, the mean call holding time is 2 min, the new call blocking probability is 1%, and the handoff call blocking probability is 0.1%. Fig. 4 plots the required number of SVC's between a base station and the AS versus cell sojourn time for different values of d . The unit of cell sojourn time is minute. It can be seen that when the mean cell sojourn time is small, which means the handoff frequency is high, the smaller the d , the less the required SVC's. This is because when the cell sojourn time is small, the number of handoff is more likely to exceed d , which means rerouting to a PVC of another base station is more likely to occur and thus leads to shorter SVC holding time. Given a cell sojourn time, the larger the d , the longer the SVC's holding time, so the required number of SVC's is larger. When the cell sojourn time is much larger than call holding time, it can be seen that for $d = 5$, the required SVC's is close to that for d is infinity. This is because almost all calls will be naturally terminated before the mobile has made more than five handoffs.

In Fig. 5, the required PVC's per cell connecting to the AS for handoff calls versus cell sojourn time is shown. It can be seen that as d increases, the required number of PVC's decreases; as the cell sojourn time increases, the required number of PVC's also decreases. This is because as the cell sojourn time increases, the handoff frequency decreases and thus the handoff call arrival rate that will request a PVC for call rerouting decreases dramatically. So although the PVC's holding time increases as cell sojourn time increases, the required number of PVC's will decrease. In our example, for the system parameter $d = 3$, it can be easily shown that when the cell sojourn time is greater than 3 s, $\lambda_h E[T_d]$ will decrease as cell sojourn time increases, but the decrease will be slower when the cell sojourn time is bigger. So the required number of PVC's per rerouting cell connecting to AS decreases as cell sojourn time increases, but the decrease is slower as cell sojourn time is bigger. Similarly, as d increases, the handoff arrival rate requiring a PVC connecting to the AS decreases while the holding time of PVC increases, the overall effect is that the required number of PVC's decreases.

In Fig. 6, the required number of PVC's per cell connecting to neighboring cells versus cell sojourn time is shown. It can be seen that as cell sojourn time increases, the number of required PVC's decreases, and, for a fixed cell sojourn time, if d decreases, the number of required PVC's would decrease, but the improvement becomes less as cell sojourn time increases. Also, when the cell sojourn time is much larger than the call holding time, for example, when cell sojourn time is 6 min and the call holding time is 2 min, the required number of PVC's

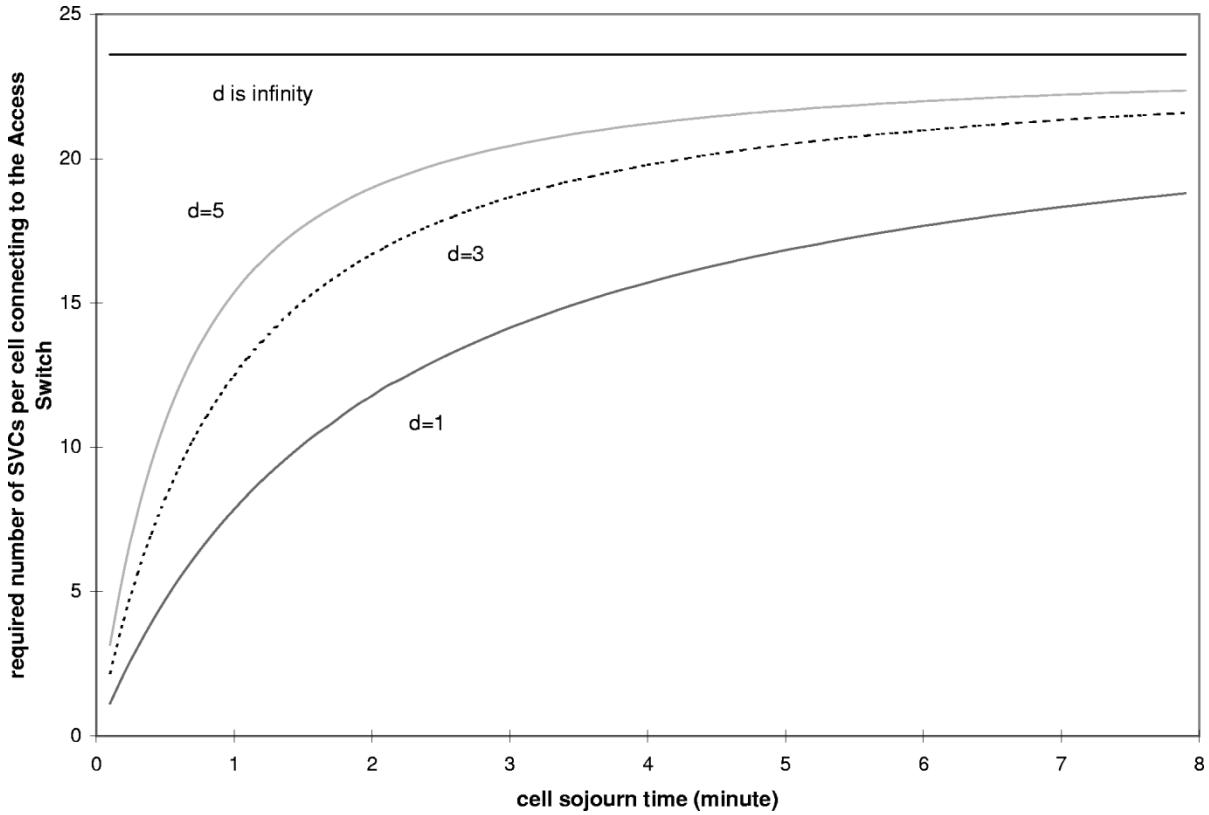


Fig. 4. The required number of SVC's for each cell connecting to an AS in the linear environment.

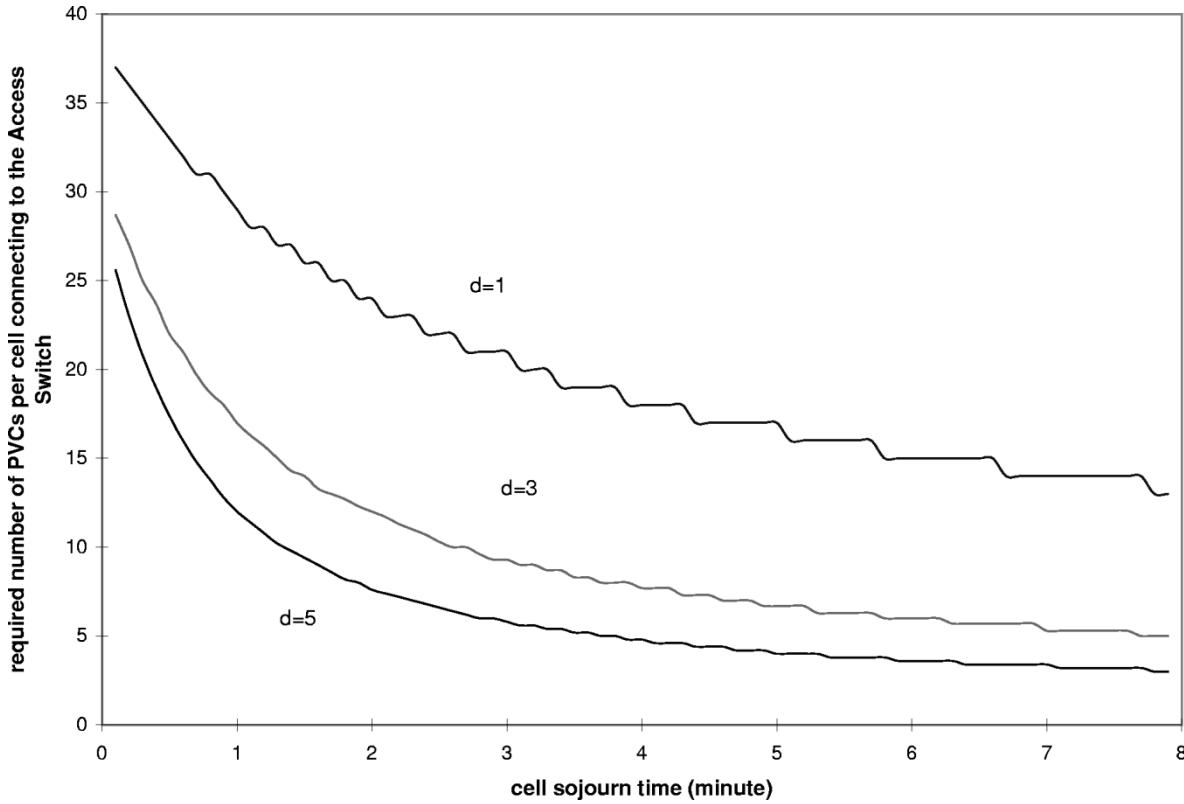


Fig. 5. The required number of PVC's per cell connecting to the AS in the linear environment.

becomes independent of d . This is because almost all calls will be naturally terminated before they needed to be rerouted.

Now we compare our scheme with the VCT [7] and PVC-based scheme [5]. Here, we consider the linear environment

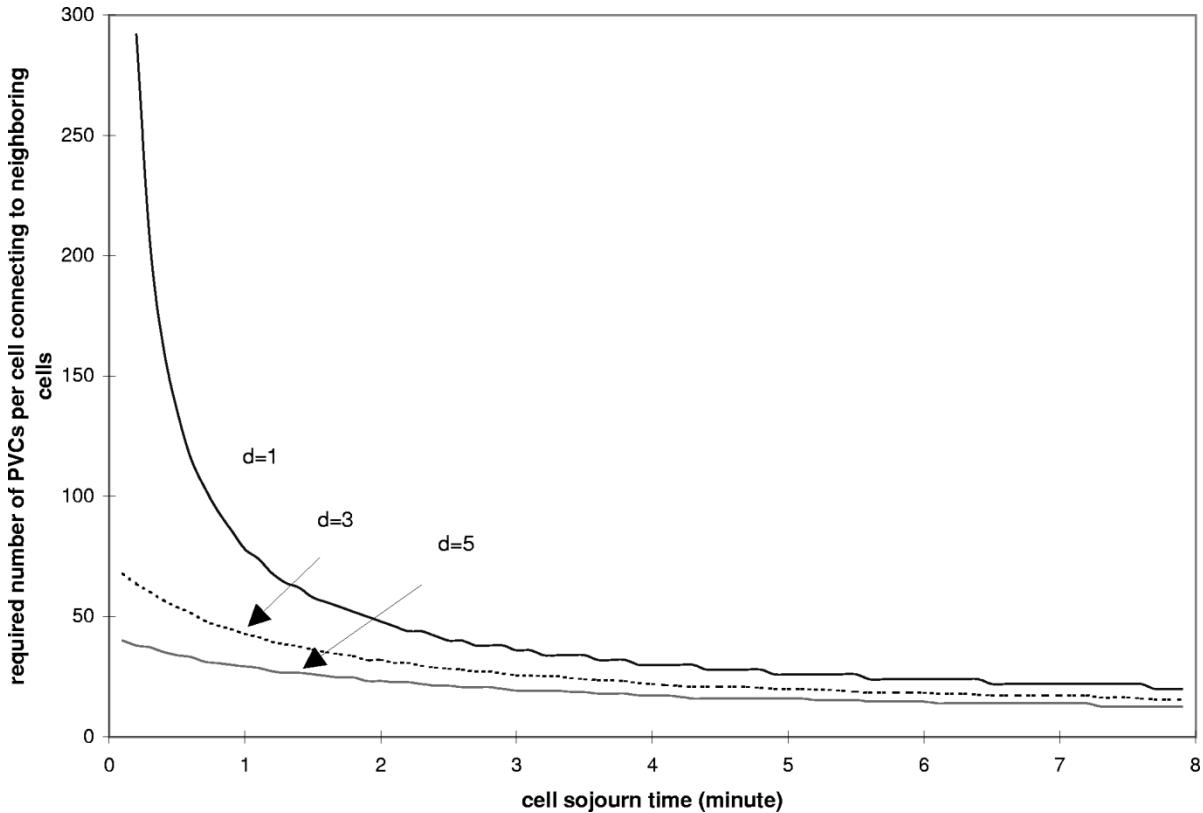


Fig. 6. The required number of PVC's per cell connecting to neighboring cells in the linear environment.

TABLE I
THE REQUIRED NUMBER OF VC'S PER
CELL FOR DIFFERENT SCHEMES

	SVC	PVC1	PVC2	total
$d=1$	11.78	24	0	35.78
$d=3$	16.68	12	23.3	51.98
$d=5$	18.99	7.6	32	58.99
$d=\infty$	23.56	0	48	71.56
VCT	1154.5	0	0	1154.5

with 49 cells. Table I shows the required number of VC's for different schemes, given the new call arrival rate is 11.9 calls per minute, the mean call holding time is 2 min, and the mean cell sojourn time is also 2 min. The new call blocking probability is 1% and handoff call blocking probability is 0.1%. In the table, the SVC column shows the number of SVC's for different schemes; the PVC1 column shows the required number of PVC's per cell connecting to the AS and the PVC2 column shows the number of required PVC's connecting to its neighboring cells. Although the VCT scheme does not need PVC's, the number of SVC's is much larger than other schemes. When $d = \infty$, our scheme becomes the PVC-based scheme as in [5] in which no PVC's connecting a cell to the AS is required, but the PVC's connecting neighboring cells is the largest. When $d = 1$, no PVC's connecting neighboring cells is required, and the total number of VC's is the smallest. However, during each handoff, the traffic path would be rerouted by AS. This means that the AS would be involved in each handoff, and hence the processing load of AS is the heaviest. So when choosing a value for d , there is a tradeoff between the traffic path efficiency and

the rerouting processing load of the AS. By choosing an appropriate system parameter d , we can keep both the processing load low and the path efficiency high. Compared with the VCT scheme, our proposed scheme requires much less total number of VC's, the bandwidth efficiency of our scheme is thus much higher.

III. PLANAR ENVIRONMENT

A. Scheme Description

In planar environment, the scheme needs some modification. As shown in Fig. 7, seven cells are grouped as a cell cluster, the central cell is defined as a rerouting cell, other six cells are called ordinary cells. Each cell belongs to one and only one cell cluster. There are PVC's connecting two neighboring cells. Each cell has SVC's connecting to the AS for new calls. Only rerouting cells have PVC's connecting to the AS for handoff calls. Here, we define a system parameter r . We also assume there is a handoff counter in each mobile. When a call is initiated in a cell, the handoff counter is set to zero. Then every time when the mobile makes a handoff, the counter is increased by one. If the counter is less than the system parameter r , the traffic path is elongated by a PVC between the current and new cells. If the counter is equal to r , the traffic path is rerouted to the new cell via the rerouting cell of the same cluster. The handoff counter is also reset to zero. If a mobile roams to a rerouting cell, the traffic path is directly rerouted to the rerouting cell, and the handoff counter is reset to zero.

Now we demonstrate how the scheme operates by an example. Referring to Fig. 7, let the base station in cell i be denoted as

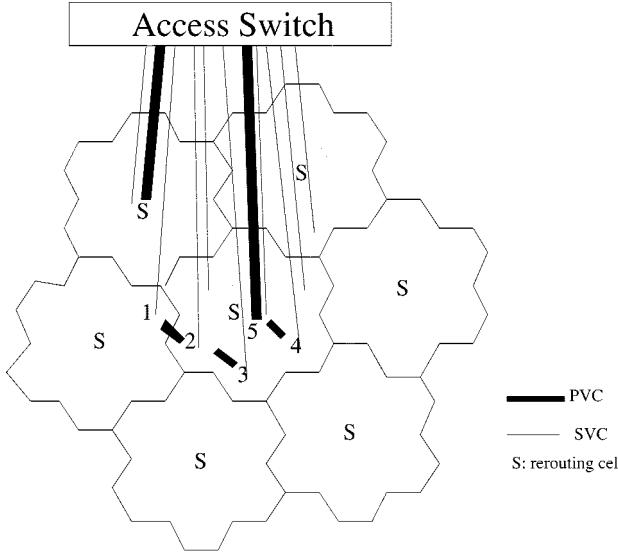


Fig. 7. The planar personal communication network.

BS_i . Assume r is three. Cells 1–4 are ordinary cells, and cell 5 is a rerouting cell. When a mobile initiates a new call in cell 1, BS_1 will establish a SVC between itself and the AS. Then, a VCI is assigned to this call, and the handoff counter is set to zero. Consider the mobile roams to its neighboring cell, say cell 2, before the call is finished. The handoff counter is increased by one. As shown by the signaling procedure depicted in Fig. 2, the mobile first sends a “handoff request” message to BS_2 , including the ID of BS_1 , the call’s VCI, QoS requirements and the value of the handoff counter. Since cell 2 is an ordinary cell, so BS_2 first checks if the value of the handoff counter is equal to r . Since it is less than r , BS_2 checks the availability of reserved PVC’s between itself and BS_1 to see whether it can support the call. If yes, it sends BS_1 a “PVC allocation” message including the identifier of the PVC to be used for path elongation. When BS_2 receives the acknowledgment for the PVC allocation from BS_1 it sends a “handoff ready” message to the mobile. Then the mobile performs the wireless channel handoff. After that, the mobile sends a “handoff” message to BS_2 indicating the completion of wireless channel handoff. After BS_2 has received the “handoff” from mobile, it sends a “handoff indication” to BS_1 . Immediately after receiving the “handoff indication” message from BS_2 and returning the acknowledgment for the indication, BS_1 routes the traffic coming from the SVC to the PVC and vice versa.

Then consider the mobile goes on to visit cell 3. The signaling procedure is the same as that of the previous handoff. Now the traffic path between the AS and the mobile terminal is: AS– BS_1 – BS_2 – BS_3 –mobile, where between AS and BS_1 is a SVC and between BS_1 and BS_2 as well as BS_2 to BS_3 are PVC’s.

Next, consider the mobile goes on to visit cell 4. The value of the handoff counter is now equal to three, which means that the traffic path should be rerouted. The handoff signaling procedure is shown in Fig. 3. The mobile first sends a “handoff request” to BS_4 . Noticing the handoff counter is equal to three, BS_4 now checks if any PVC connecting to rerouting cell 5 is available. If a PVC is available, BS_4 will send a PVC allocation and rerouting request message to rerouting cell 5. When BS_5 receives this mes-

sage from BS_4 , BS_5 will check if any PVC between BS_5 and the AS is available. If that is true, BS_5 will send a “PVC allocation” message to the AS including the call’s VCI and the identifier of the PVC to be used for the path rerouting. When the AS has received the “PVC allocation” from BS_5 , it returns an acknowledgment to BS_5 , which then also returns an ACK to BS_4 , indicating that the path to the AS for rerouting is ready. BS_4 now sends “handoff ready” to the mobile. After receiving the “handoff ready” from BS_4 , the mobile performs the wireless channel handoff. When that is finished, the mobile sends “handoff” to BS_4 . Meanwhile the mobile also sends a “PVC release” message to BS_3 . After receiving the “handoff” from the mobile, BS_4 sends “handoff indication” to BS_5 . After BS_5 has received this message, it updates its routing table and sends a “handoff indication” to the AS. After the AS has received the “handoff indication,” it updates its routing table, and returns an ACK to that indication. Upon receiving the ACK from the AS, BS_5 also returns an ACK to BS_4 . At the same time, when BS_3 has received the “PVC release” from the mobile, it releases the reserved PVC between itself and BS_2 and sends a similar message to BS_2 . BS_2 releases the reserved PVC and propagates the message to BS_1 . Then BS_1 will release the SVC connecting to the AS for the call.

B. Performance Analysis

Making the same assumptions as in the linear case, a VC connecting the cell’s base station to the AS or another cell’s base station is occupied by the mobile. The VC can be released in three cases: 1) the connection is naturally terminated; 2) the connection is forced to be terminated due to handoff blocking; and 3) the traffic path is rerouted at its i th handoff. As in the linear case, the mean of the VC’s holding time is

$$E[T_i] = \frac{1 - \left[\frac{\mu_R(1 - p_f)}{\mu_M + \mu_R} \right]^i}{\mu_M + p_f \mu_R}.$$

Let us define the following variables first:

- N_1 required number of PVC’s connecting an ordinary cell to one of its ordinary neighboring cells;
- N_2 required number of PVC’s connecting an ordinary cell to the rerouting cell of the same cluster;
- N_3 required number of PVC’s connecting a rerouting cell to the AS;
- N_4 required number of SVC’s connecting a cell to the AS;
- λ_{h1} arrival rate of handoff calls coming from an ordinary cell to an ordinary cell that require a PVC connecting the two cells;
- λ_{h2} arrival rate of handoff calls coming to an ordinary cell that requires a PVC connecting the ordinary cell to the rerouting cell of the same cluster;
- λ_{h3} arrival rate of handoff calls requiring a PVC connecting the rerouting cell to the AS.

Now let us derive the handoff call arrival rate. Let p_i be the probability that a call will make at least i handoff requests. Then we have

$$p_i = (1 - p_f)^{i-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^i.$$

The total handoff arrival rate to a cell λ_h is

$$\begin{aligned}\lambda_h &= \lambda_0(1-p_0) \sum_{i=1}^{\infty} (1-p_f)^{i-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^i \\ &= \frac{\lambda_0 \mu_R (1-p_0)}{\mu_R p_f + \mu_M}.\end{aligned}$$

We divide the handoff calls into two types: elongating handoff requiring a PVC connecting the current and new cells to elongate the traffic path and rerouting handoff requiring the call to be rerouted to the new cell via the rerouting cell in the same cluster. We further divide the rerouting handoff into two types: early rerouting handoff when the mobile enters a rerouting cell while its handoff counter is less than the system parameter r and normal rerouting handoff when the handoff counter is equal to the system parameter r . Let us define q_i as the probability that a mobile makes exactly i handoffs during its call. From the Appendix, q_i is given by

$$q_i = (1-p_f)^{i-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^i \left[1 - \frac{\mu_R(1-p_f)}{\mu_R + \mu_M} \right].$$

Consider a mobile initiating a call in a cell. The mean number of handoffs n the mobile may make during its call is given by

$$\begin{aligned}n &= \sum_{i=1}^{\infty} i q_i \\ &= \sum_{i=1}^{\infty} (1-p_f)^{i-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^i \\ &= \frac{\mu_R}{\mu_R p_f + \mu_M}.\end{aligned}$$

Consider the mobile has never been a rerouting cell during its call, which means that there is no early rerouting handoff. Denote, for this case, the mean number of normal rerouting hand-

offs be n_1 and the mean number of elongating handoffs be n_2 . Then, the r th, $(2r)$ th, \dots , (kr) th handoff (k is a positive integer) are normal rerouting handoffs. So

$$\begin{aligned}n_1 &= \sum_{i=1}^{\infty} (1-p_f)^{ir-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^{ir} \\ &= \frac{(1-p_f)^{r-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}{1 - (1-p_f)^r \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}\end{aligned}$$

and

$$\begin{aligned}n_2 &= \sum_{i=1}^{\infty} (1-p_f)^{i-1} \left(\frac{\mu_R}{\mu_R p_f + \mu_M} \right)^i - n_1 \\ &= \left(\frac{\mu_R}{\mu_R p_f + \mu_M} - \frac{(1-p_f)^{r-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}{1 - (1-p_f)^r \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r} \right).\end{aligned}$$

Now, consider the mobile was in a rerouting cell only at its $(kr+i)$ th handoff where k and i are positive integers and $1 \leq i \leq r-1$. For this case, denote the mean number of normal rerouting handoffs be n'_1 and the mean number of elongating handoffs be n'_2 . It can be seen the r th, $(2r)$ th, \dots , (kr) th handoffs and $(kr+i+r)$ th, $(kr+i+2r)$ th, \dots handoffs are normal rerouting handoffs. So we have (5a), given at the bottom of the page.

It can be seen that $n'_1 \leq n_1$ and $n'_2 \leq n_2$. Assume a mobile has experienced two early rerouting during a call, and the mean of the number of normal rerouting the mobile makes during this call be n''_1 . It can be easily proved that $n''_1 \leq n_1$. So no matter how many times of early rerouting a mobile may experience during a call, n_1 is always the upper bound of the mean of the number of normal rerouting the mobile may make during a call,

$$\begin{aligned}n'_1 &= \sum_{j=1}^k (1-p_f)^{jr-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^{jr} + \sum_{l=1}^{\infty} (1-p_f)^{kr+i+lr-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^{kr+i+lr} \\ &= \frac{(1-p_f)^{r-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r \left[1 - \left(\frac{(1-p_f)\mu_R}{\mu_R + \mu_M} \right)^{kr} + \left(\frac{(1-p_f)\mu_R}{\mu_R + \mu_M} \right)^{kr+i} \right]}{1 - (1-p_f)^r \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}\end{aligned}$$

and

$$\begin{aligned}n'_2 &= \frac{\mu_R}{\mu_R p_f + \mu_M} - n'_1 - (1-p_f)^{kr+i-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^{kr+i} \\ &= \frac{\mu_R}{\mu_R p_f + \mu_M} - (1-p_f)^{kr+i-1} \left(\frac{\mu_R}{\mu_M + \mu_R} \right)^{kr+i} \\ &\quad - \frac{(1-p_f)^{r-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r \left[1 - \left(\frac{(1-p_f)\mu_R}{\mu_R + \mu_M} \right)^{kr} + \left(\frac{(1-p_f)\mu_R}{\mu_R + \mu_M} \right)^{kr+i} \right]}{1 - (1-p_f)^r \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}\end{aligned} \quad (5a)$$

and it is used as the estimation of the mean of the number of normal rerouting a mobile may make during a call.

Consider a cell cluster as a unit. The mean rate of normal rerouting handoffs and elongating handoffs due to all calls initiated in the cluster are $7\lambda_0(1 - p_0)n'_1$ and $7\lambda_0(1 - p_0)n'_2$, respectively. Now consider the arrival rate of handoffs coming to the cluster. Define λ'_1 and λ'_2 as the mean arrival rates of normal rerouting handoffs and elongating handoffs due to an ordinary neighboring cell, respectively. The mean arrival rate of normal rerouting handoffs to the cluster is $30\lambda'_1 + 6\lambda'_1$, where the first term is the rate coming to the six ordinary cells, and the second term is that coming to the rerouting cell. If the system is stable, the arrival rate of normal rerouting handoffs to the cluster should be equal to that generated by those calls initiated within the cluster. So, we have

$$\begin{aligned} 36\lambda'_1 &= 7\lambda_0(1 - p_0)n'_1 \\ &\leq 7\lambda_0(1 - p_0)n_1 \end{aligned}$$

or

$$\lambda'_1 \leq \frac{7}{36}\lambda_0(1 - p_0)n_1$$

which will be used as the upper bound of λ'_1 .

Similarly, the mean of total arrival rate of elongating handoffs in a cluster is $\lambda_h + 30\lambda'_2$, where the first term is elongating handoffs from the rerouting cell to its neighboring cells, and the second term is elongating handoffs from ordinary cells to ordinary cells. Again, we have

$$\begin{aligned} 30\lambda'_2 + \lambda_h &= 7\lambda_0(1 - p_0)n'_2 \\ &\leq 7\lambda_0(1 - p_0)n_2 \end{aligned}$$

or

$$\lambda'_2 \leq \frac{1}{5}\lambda_h - \frac{7}{30}\lambda_0(1 - p_0) \frac{(1 - p_0)^{r-1} \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}{1 - (1 - p_0)^r \left(\frac{\mu_R}{\mu_R + \mu_M} \right)^r}$$

which will be used as the upper bound of λ'_2 .

According to the definition, we have

$$\lambda_{h1} = \lambda'_2.$$

Now consider the call arrival rate of handoffs that request PVC's connecting a ordinary cell to the rerouting cells. It consists of two components: handoff calls coming from the rerouting cell with mean arrival rate $\lambda_h/6$; rerouting handoffs from ordinary neighboring cells with mean arrival rate $5\lambda'_1$. So, we have

$$\lambda_{h2} = \lambda_h/6 + 5\lambda'_1. \quad (6)$$

For the mean arrival rate of handoffs requesting a PVC connecting the rerouting cell to the ATM switch, it consists of the rate of handoffs into the rerouting cells and the rate of rerouting handoffs into the six ordinary cells surrounding the special cells. Therefore, the upper bound of the arrival rate of handoffs re-

questing a PVC connecting the rerouting cell to the AS is given by

$$\lambda_{h3} = \lambda_h + 30\lambda'_1. \quad (7)$$

By this scheme, an SVC can be occupied at most $E[T_r]$ if the mobile does not meet a rerouting cell during its first r handoffs, or $E[T_i]$ if it meets a rerouting cell at its i th handoff. For simplicity, here we give an upper bound for the required number of SVC's, N_s . We have

$$N_s = \lambda_0(1 - p_0)E[T_r].$$

Now consider an ordinary cell. We will derive the required number of PVC's connecting to an ordinary neighboring cell. When a handoff call comes to the ordinary cell from the considered neighboring cell and its handoff counter is only one, according to the scheme, the longest mean holding time of the PVC for elongation is $E[T_{r-1}]$; if the handoff counter is two, the longest holding time is $E[T_{r-2}]$ and so on. Here, we give an upper bound, assuming all the PVC's holding time is $E[T_{r-1}]$. We can model that as a $M/M/m/m$ queuing system, where the arrival rate is λ_{h1} and the average holding time is $E[T_{r-1}]$, so we have

$$p_f = \frac{(\lambda_{h1}E[T_{r-1}])^{N_1}/N_1!}{\sum_{n=0}^{N_1} (\lambda_{h1}E[T_{r-1}])^n/n!}. \quad (8)$$

By (8), we can obtain the upper bound of N_1 that satisfies the handoff call blocking probability p_f .

Next, we will derive N_2 the number of PVC's connecting an ordinary cell to the rerouting cell in the same cell cluster. If the handoff call is from the rerouting cell, the upper bound of the PVC's mean holding time is $E[T_{r-1}]$; if the call is from another ordinary cell, the mean holding time is $E[T_r]$. As before, we can model it as a $M/M/m/m$ queuing system with arrival rate λ_{h2} and mean holding time $E[T_r]$. Then

$$p_f = \frac{(\lambda_{h2}E[T_r])^{N_2}/N_2!}{\sum_{n=0}^{N_2} (\lambda_{h2}E[T_r])^n/n!}. \quad (9)$$

For the required number of PVC's connecting a rerouting cell to the AS, we model it as a $M/M/m/m$ system with call arrival rate λ_{h3} and mean holding time $E[T_r]$, we have

$$p_f = \frac{(\lambda_{h3}E[T_r])^{N_3}/N_3!}{\sum_{n=0}^{N_3} (\lambda_{h3}E[T_r])^n/n!}. \quad (10)$$

Finally, we consider the processing load defined in previous section. In the planar environment, consider a cell cluster. The total arrival rate of handoffs coming to any one of cells in the cluster requiring path rerouting is λ_{h3} , while the total arrival rate of handoffs coming to any one of the cells in the cluster is $7\lambda_h$. So the processing load in the planar environment is

$$p_l = \frac{\lambda_{h3}}{7\lambda_h}. \quad (11)$$

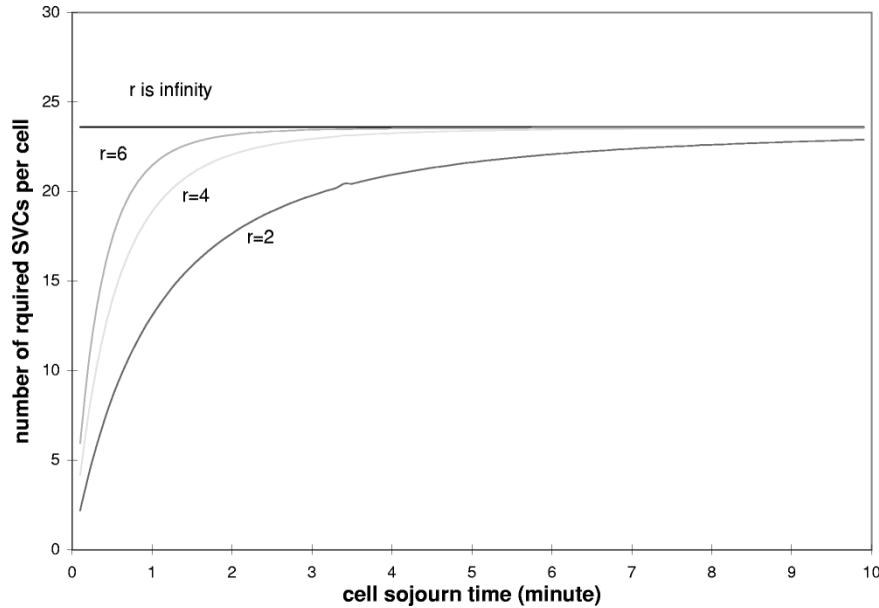


Fig. 8. The required number of SVC's per cell connecting to the AS in the planar environment.

C. Numerical Results

In this section, some numerical results are presented, given the new call arrival rate is 11.9 calls per minute per cell, the mean call holding time is 2 min, the new call blocking probability is 1% and the handoff call blocking probability is 0.1%. Fig. 8 plots the required number of SVC's between a base station and the AS in the planar environment versus cell sojourn time for different values of r . The unit of cell sojourn time is minute. It can be seen that when the mean cell sojourn time is small, which means the handoff frequency is high, the smaller the r , the less the required SVC's. This is because when the cell sojourn time is small, the number of handoff is more likely to exceed r , which means rerouting is more likely to occur and thus leads to shorter SVC holding time. Given a cell sojourn time, the larger the r , the longer the SVC's holding time, so the required number of SVC's is larger. When the cell sojourn time is much larger than call holding time, it can be seen that for $r = 6$, the required SVC's is the same as that for r is infinity. This is because almost all calls will be naturally terminated before the mobile has made more than six handoffs.

In Fig. 9, the required PVC's in an ordinary cell connecting to one of its ordinary neighboring cell versus cell sojourn time is shown. It can be seen that as r increases, the required number of PVC's increases; as the cell sojourn time increases, the required number of PVC's decreases. This is because as the cell sojourn time increases, the handoff frequency decreases and thus the arrival rate of elongating handoff decreases dramatically. So even though the PVC's holding time increases with cell sojourn time, the required number of PVC's will decrease. Similarly, as r increases, the arrival rate of elongating handoff increases and the holding time of PVC increases, so the required number of PVC's increases.

In Fig. 10, the required number of PVC's of each ordinary cell connecting to the rerouting cell in the same cluster versus cell sojourn time is shown. It can be seen that as cell sojourn time

increases, the number of required PVC's decreases. Again, that is because as cell sojourn time increases, the handoff call arrival rate decreases rapidly, so even though the VC's holding time increases, the required PVC's still decreases. When cell sojourn time is very small, the required number of PVC's increases with r . This is because, for small cell sojourn time, the increase in holding time due to increase in r becomes more dominant. On the other hand, when cell sojourn time is large, the required number of PVC's decreases as r increases. This is because, for large cell sojourn time, the increase in handoff arrival rate due to decrease in r is more dominant.

Fig. 11 shows the required number of PVC's per rerouting cell connecting to the AS versus the cell sojourn time for various system parameter r . The PVC's are used for two types of handoff calls: the first type is those handoff calls coming to the rerouting cell; the other type is those handoff calls coming to the ordinary cells in the same cluster and requiring path rerouting. It can be seen that, as cell sojourn time increases, the required number of PVC's decreases. Again, this is because the handoff call arrival rate decreases more rapidly than the rate of increase in VC holding time. It can also be seen that when cell sojourn time is very small, as r increases, the required number of PVC's increases; for a large cell sojourn time (greater than 0.5 min), the larger the r , the smaller the number of required PVC's. This is because, for small cell sojourn time, as r increases, the VC's holding time increases rapidly. So even though the arrival rate of the second type of handoff calls decreases, the total required PVC's still increases. For large cell sojourn time, the effect of r on VC's holding time is much smaller, so as the arrival rate of the second type of handoff call decreases, the required number of PVC's decreases. In our example, it can be shown that when cell sojourn time is less than 20 s, $\lambda_{h3}E[T_r]$ will increase as r increases; when cell sojourn time is greater than 20 s, $\lambda_{h3}E[T_r]$ will decrease as r increases.

Fig. 12 shows the processing load versus cell sojourn time for different r . It can be seen that as cell sojourn time increases,

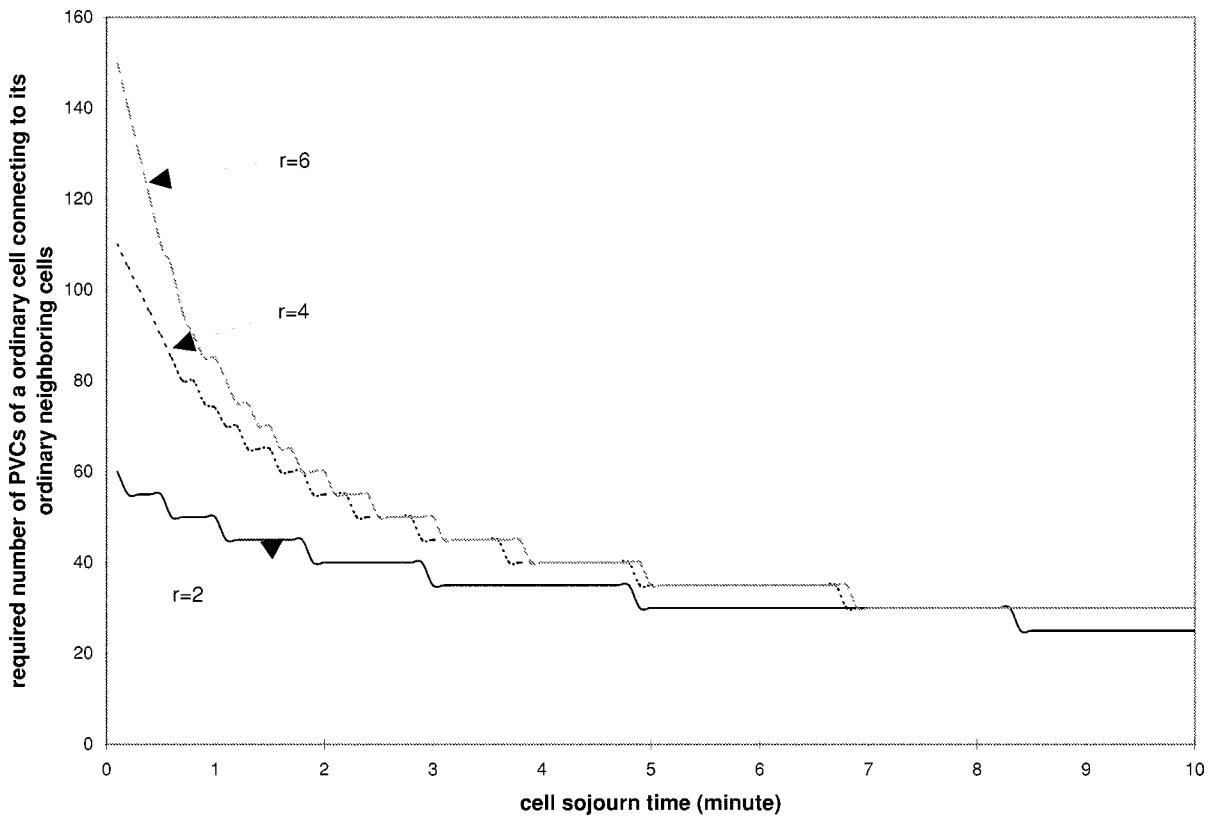


Fig. 9. The required number of PVC's for an ordinary cell connecting to an ordinary neighboring cell in planar environment.

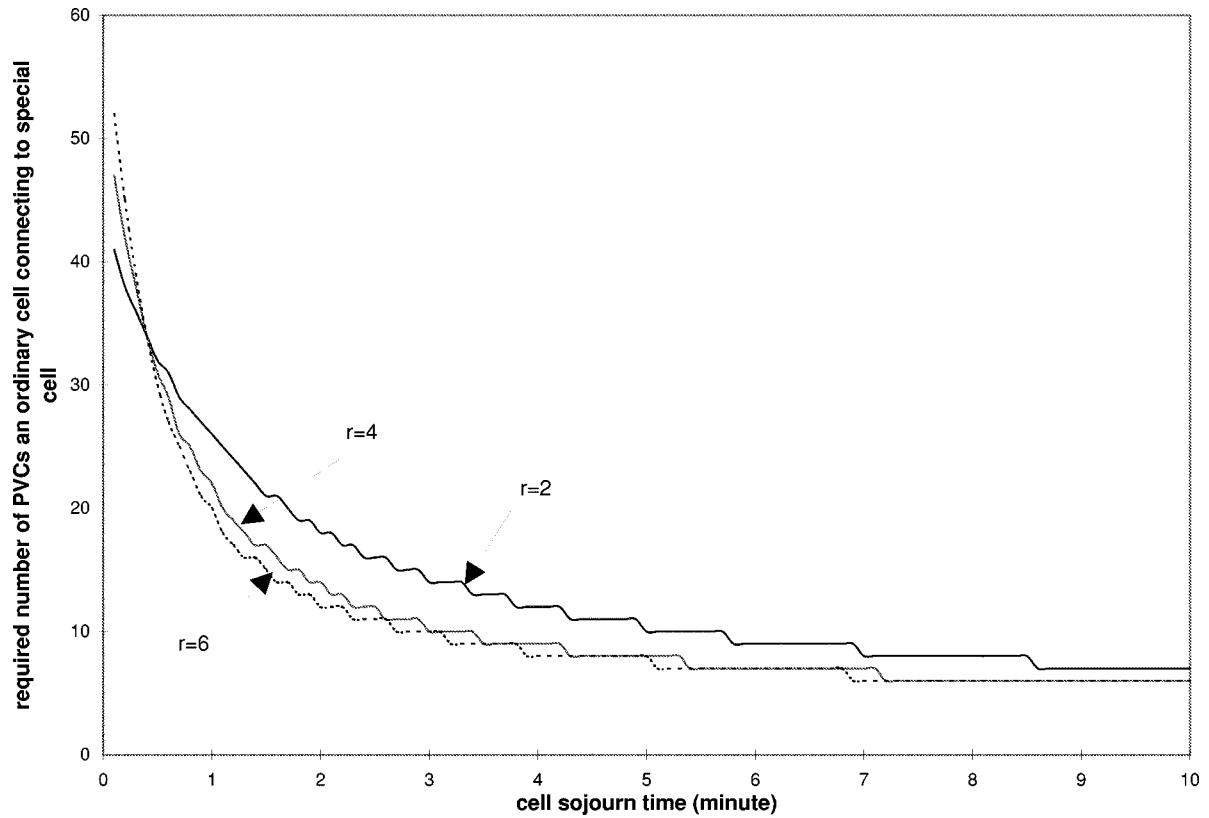


Fig. 10. The required number of PVC's for an ordinary cell connecting to the rerouting cell in the same cell cluster.

the processing load decreases. It is because, given a fixed r , as cell sojourn time increases, the handoff calls that need to

be rerouted decreases, so the processing load decreases. It can also be seen that, given a fixed cell sojourn time, the processing

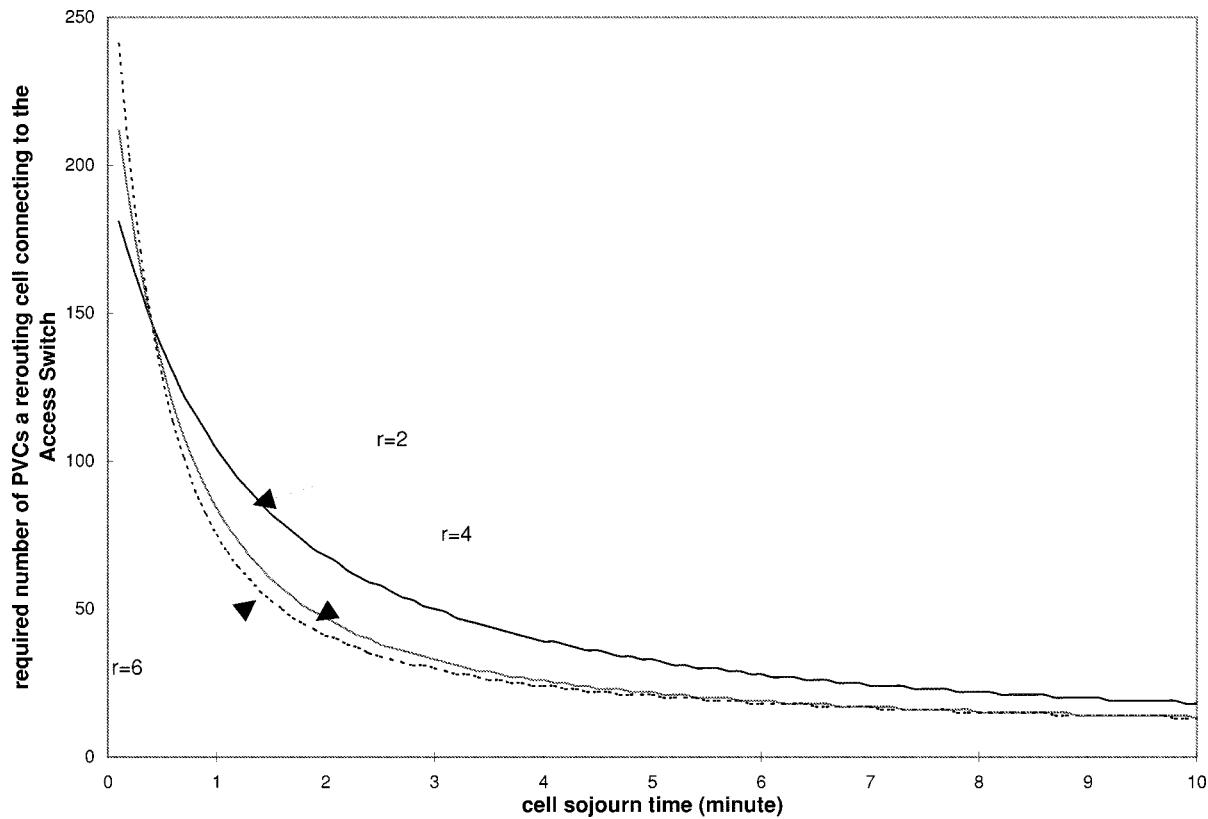


Fig. 11. The required number of PVC's for each rerouting cell connecting to the AS in the planar environment.

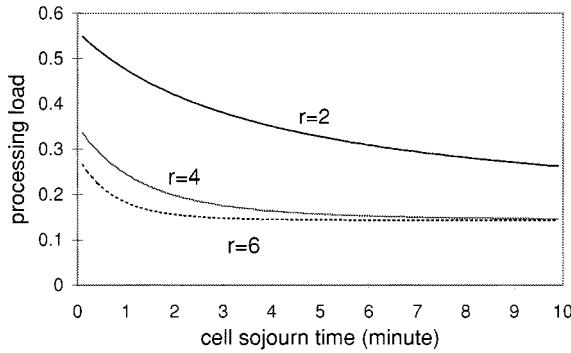


Fig. 12. The processing load versus cell sojourn time.

load decreases as r increases. That is because, as r increases, the proportion of handoffs that need to be rerouted decreases, so processing load decreases. So to a fixed cell sojourn time, we can choose parameter r to keep the processing load light.

Finally, we again compare our schemes to the VCT and PVC-based schemes. Assume there are 49 cells in a planar environment. Other parameters are the same as that in the linear environment. Table II shows the results. It can be seen that our scheme requires less VC's than the other two schemes.

IV. GENERALIZED SCHEME WITH MULTIPLE AS's

So far, we have described our scheme in the situation that all cells are connected to the same AS. In general, an AS only covers certain geographical area and hence cells in different

TABLE II
THE REQUIRED NUMBER OF VC's FOR DIFFERENT SCHEMES

	SVC	PVC1	PVC2	total
$r=2$	17.66	9.7	49.7	77.06
$r=4$	22.07	6.7	57.4	86.17
$r=6$	23.17	5.8	61.7	90.67
$r=\infty$	23.56	0	72	97.56
VCT	1154.5	0	0	1154.5

areas are connected to different AS's. Fig. 13 shows the architecture with multiple AS's in linear environment where cells 0–23 are connected to AS 1; cells 23–47 are connected to AS 2. We denote the area served by the same AS as switch area. In our example, cells 0–23 belong to switch area 1; cells 24–47 belong to switch area 2. Also, we denote the cells on the boundary of switch area as boundary cells. Here, we explain how our scheme handles handoffs across neighboring switch areas.

The rerouting cells are assigned as follows: for a switch area, the most left cell is assigned as a rerouting cell. For a given system parameter d , after every d cells, another rerouting cell is assigned. The most right cell (the right boundary cell) is also assigned as a rerouting cell. So for a switch area with n cells, there are $m = \lceil (n + d - 1)/d \rceil$ rerouting cells. In our example, we assume $d = 4$. Consider the switch area 1 with 24 cells. We label cells from left to right as cell 0, cell 1, and so on. Cells 0, 4, ..., 20, 23 are rerouting cells. If a mobile roams within a switch area and never comes to the boundary cells, the operation is the same as that described before. When a mobile comes to a boundary cell, a signal is sent to find out the nearest common node (NCN) between the AS's of the current and new

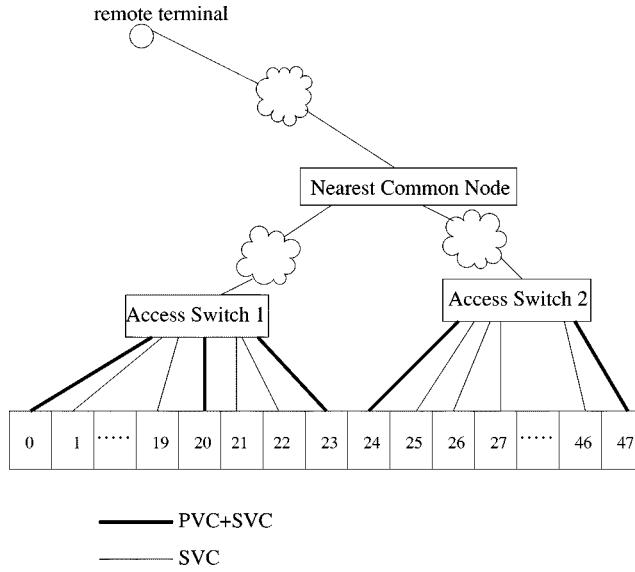


Fig. 13. The general scenario with multiple AS's in the linear environment.

switch areas, as explained in details in [6], and then set up a SVC between the NCN and the AS of the new switch area. If the mobile goes on to the new switch area, the existing traffic path between the NCN and the mobile is forced to be rerouted to the path consisting of the newly established SVC and a PVC between the AS in the new switch area and the mobile.

This is illustrated by the following example. Assume a mobile initiates a call in cell 21. The connection path to the remote terminal is: mobile- BS_{21} -AS 1-NCN-remote terminal. If the mobile roams to cell 22, the path is simply elongated by a PVC between BS_{22} and BS_{21} . If the mobile handoffs to cell 23 successfully, as cell 23 is a boundary cell (its right neighboring cell 24 is connected to AS 2), BS_{23} will send a signal to the remote terminal to find out the NCN between AS 1 and AS 2, and set up a SVC between the AS 2 and NCN for the call. Now if the mobile roams to cell 24, the traffic path is rerouted. A PVC between AS 2 and BS_{24} is assigned for the call. The traffic path between the mobile and remote terminal becomes: mobile- BS_{24} -AS 2-NCN-remote terminal.

Fig. 14 shows the architecture in planar environment with multiple AS's. Again we denote the area served by the same AS as switch area. For example, cells 1-4 belong to switch area 1; cells 5-7 belong to switch area 2. We also denote the cells on the boundary of switch area as boundary cells. Here, we do not assign all boundary cells as rerouting cells.

If a mobile roams within a switch area and never comes to the boundary cells, the operation is the same as that described in the previous section. When the mobile handoffs from one switch area to another, the operation is a bit different from that in linear case. When it comes to a boundary cell, a signal is sent to find out the nearest common node (NCN) between the AS's of the current and new switch areas, and then set up a SVC between the NCN and the AS of the new switch area. If the mobile goes on to the new switch area, the existing traffic path between the NCN and the mobile is forced to be rerouted. If the newly entered cell is a rerouting cell, the new path only consists of the newly established SVC and a PVC between the AS of the new switch area and the base station of the new cell. If the new cell is an

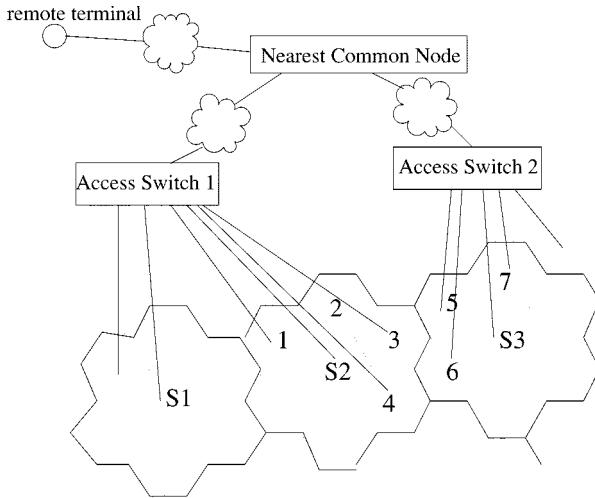


Fig. 14. The general scenario with multiple AS's.

ordinary cell, the new traffic path includes the new SVC, a PVC between the new AS and the rerouting cell of the cell cluster to which the new cell belongs, and a PVC from the rerouting cell to the new cell.

For example, a mobile roams in switch area 1. When it enters cell 3, as cell 3 is a boundary cell (its neighboring cells 5 and 6 are connected to AS 2), BS_3 will send a signal to the remote terminal to find out the NCN between AS 1 and AS 2, and set up a SVC between the AS 2 and NCN for the call. Note that we do not set up the SVC to base stations because there are more than one possible cells that the mobile may enter.

Now if the mobile roams to cell 5, the traffic path is rerouted through rerouting cell $S3$. A PVC between AS 2 and cell $S3$ and a PVC between BS_5 and BS_{S3} are assigned for the call. The traffic path between the mobile and remote terminal becomes: mobile- BS_5 - BS_{S3} -AS 2-NCN-remote terminal.

Now we investigate the impact of the size of switch area to the performance of our scheme. We analyze the performance of our scheme in linear environment. Referring to Fig. 13, here we assume the sizes of all switch areas are the same, each switch area includes n cells, for the simplicity of comparison. And we assume the number of cells in the system is infinity. There are two types of processing by the network in the multiple AS environment, one for rerouting the traffic path by the AS and another for searching the NCN when the mobile is ready to roam from one switch area to another. Here, we define two processing measures: p_c for the processing load of searching the NCN, and p_r for the traffic path rerouting by the AS

$$p_c = \frac{\text{number of handoffs requiring NCN search}}{\text{total number of handoffs}}$$

$$p_r = \frac{\text{number of handoffs requiring traffic path rerouting}}{\text{total number of handoffs}}.$$

The following variables are also defined:

n_{si} average of occupied number of SVC's connecting the base station in cell i to its AS;

n_{pi}^l number of required PVC's connecting base station in cell i to its left neighboring cell;

n_{pi}^r	number of required PVC's connecting the base station of cell i to its right neighboring cell;
n_{pi}^a	number of required PVC's connecting the base station of cell i to its AS;
N_s	average of occupied number of SVC's by mobiles in one cell connecting the base station of the cell to AS;
N_{p1}	average of required number of PVC's connecting a rerouting cell to the AS per cell;
N_{p2}	average of required number of PVC's connecting on e cell to its neighboring cells per cell.

Now we derive the expressions for these variables. Consider a switch area with n cells. The mean handoff call arrival rate of a cell is λ_h , as explained before. We assume that a handoff call is from the left or right neighboring cell with the same probability. Consider the right boundary cell of a considered switch area, which is cell 23 of the switch area 1 in Fig. 13. Only the handoff calls from the left neighboring cell need to search the NCN between the current and new AS's. For the handoff calls from the right neighboring cell, for example, handoff calls from cell 24 to cell 23, as when the mobile is in the previous cell, cell 24, the NCN has been found out between the two AS's, it is not necessary to perform that again when the mobile comes to the considered cell. There is a similar scenario to the left boundary cell. So the mean handoff calls of the switch area requiring NCN search is only $2 \times \lambda_h/2$. The NCN search processing load is

$$p_c = \frac{\lambda_h}{n\lambda_h} = \frac{1}{n}. \quad (12)$$

Given the system parameter d , there are $m = \lceil (n+d-1)/d \rceil$ rerouting cells for a switch area with n cells. So the rerouting processing load is

$$p_r = \frac{m\lambda_h}{n\lambda_h} = \frac{m}{n}. \quad (13)$$

The derivations for N_s , N_{p1} , and N_{p2} are a bit more complicated. Here, we give an example to show how to do it. Referring to Fig. 13, assume $n = 24$ and $d = 4$. The VC requirements for cells 1–19 are the same as cells with one AS case. So only the requirements for cells 0, 20, 21, 22, and 23 need to be derived. As an example, we derive the VC requirements for cell 22.

If a mobile in cell 22 initiates a call, a SVC is captured for this call. Then if the mobile moves left, as cell 20 is a rerouting cell, the holding time for this SVC is T_2 ; if the mobile moves right, again as cell 23 is a rerouting cell, the holding time for this SVC is T_1 . So the average captured SVC's for cell 22 is

$$n_{s22} = \lambda_0(1 - p_0)(E[T_2] + E[T_1])/2$$

where λ_0 is the mean new call arrival rate and p_0 is the new call blocking probability.

As cell 22 is an ordinary cell, $n_{p22}^a = 0$. We now derive n_{p22}^l . When a call handoffs from cell 21 to cell 22, a PVC connecting cell 22 to cell 21 is captured by this call. Then the mobile may

TABLE III
COMPARISON OF DIFFERENT SWITCH AREA SIZES

switch area	p_c	p_r	SVC	PVC1	PVC2	total
s=3	0.333	0.667	9.79	28.67	12	50.46
s=7	0.143	0.428	12.24	22.43	21.14	55.81
s=14	0.07	0.36	13.23	18.36	31.71	63.3
s=24	0.04	0.29	13.67	16.37	34.5	64.54
s=∞	0	0.25	14.1	14.25	37	65.35

move on to cell 23. As cell 23 is a rerouting cell, so the holding time for this PVC is T_1 . Similar to the derivation before, we have

$$p_f = \frac{(\lambda_h E[T_1])^{n_{p22}^l} / n_{p22}^l!}{\sum_{j=0}^{n_{p22}^l} (\lambda_h E[T_1])^j / j!}. \quad (16)$$

When a call handoffs from cell 23 to cell 22, a PVC connecting cell 22 to cell 23 is seized by this call. Then the mobile may move on to cell 21. As cell 20 is a rerouting cell, so the holding time for this PVC is T_2 . Then we have

$$p_f = \frac{(\lambda_h E[T_2])^{n_{p22}^r} / n_{p22}^r!}{\sum_{j=0}^{n_{p22}^r} (\lambda_h E[T_2])^j / j!}. \quad (17)$$

Now we have

$$N_s = \frac{\sum_{i=0}^{n-1} n_{si}}{n} \quad (18)$$

$$N_{p1} = \frac{\sum_{i=0}^{n-1} n_{pia}}{n} \quad (19)$$

$$N_{p2} = \frac{\sum_{i=0}^{n-1} (n_{pi}^l + n_{pi}^r)}{n}. \quad (20)$$

Table III shows the numerical results, given the new call arrival rate is 11.9 calls per minute per cell, the mean call holding time is 2 min, the new call blocking probability is 1%, the handoff call blocking probability is 0.1%, and the mean cell sojourn time is 1 min. It can be seen that the processing load for NCN search is only affected by the size of switch area. The main trend of processing load for path rerouting decreases as the size of switch area increases. A similar scenario can also be observed in VC requirement. When the size of switch area is less than the system parameter d , the performance is only affected by the size of switch area. When the size of switch area is larger, the processing load for path rerouting and VC requirements are affected both by the switch area size and system parameter d . So when switch area size is large, we can appropriately choose the system parameter d to make a tradeoff between the processing load low and the path efficiency high. Similar scenarios can also be observed in planar environment. That is,

when the switch area size is large, by carefully choosing the system parameter r , we can keep both the processing load low and path efficiency high.

V. CONCLUSIONS

In this paper, an efficient handoff scheme for wireless ATM networks is proposed. In this scheme, all cells are connected to its neighboring cells and the AS by PVC's and SVC's, respectively. PVC's are used for handoff calls, while SVC's are used by new calls. Rerouting cells are evenly spaced in the network. If a mobile roams to a neighboring ordinary cell, its traffic path is simply elongated by the PVC connecting the old and new cells. If a mobile roams to a rerouting cell, its traffic path is rerouted to the rerouting cell. By using PVC's for handoff calls, we can guarantee the handoff process to be fast and seamless. Also, by path rerouting, we can dramatically improve the path efficiency. Suitably selecting the hop limiting parameter r allows most handoff calls to be handled by the involved base stations using PVC's, the added processing load to the network due to frequent handoffs can therefore be kept low. An analytical model has been developed to dimension the required VC's under our scheme. It has been shown that our scheme requires less VC's than the VCT and connection extension schemes.

APPENDIX

Here, we derive the expression of q_i , the probability that a mobile makes exactly i handoffs during its call. Consider a mobile initiates a call in a cell. The call's holding time is exponentially distributed with mean $1/\mu_M$. The cell sojourn time is also exponentially distributed with mean $1/\mu_R$. If the call's holding time is greater than the cell sojourn time, the mobile will move out of the cell, that is, a handoff attempt will be made. Define this probability as p_h , it can be easily shown that

$$p_h = \frac{\mu_R}{\mu_M + \mu_R}.$$

Assume the handoff call blocking probability be p_f . Now we derive the probability of q_1 . It consists of two cases: 1) the user's first handoff request is blocked and 2) the first handoff is successful, and the call is naturally terminated in the second cell. The probability for the first case is $(\mu_R/(\mu_R + \mu_M))p_f$; the probability for the second case is $(\mu_R/(\mu_R + \mu_M))(1-p_f)(1-(\mu_R/\mu_R + \mu_M))$. So we have

$$\begin{aligned} q_1 &= \frac{\mu_R}{\mu_M + \mu_R}p_f + \frac{\mu_R}{\mu_R + \mu_M}(1-p_f)\left[1 - \frac{\mu_R}{\mu_R + \mu_M}\right] \\ &= \frac{\mu_R}{\mu_R + \mu_M}\left[1 - \frac{\mu_R(1-p_f)}{\mu_R + \mu_M}\right]. \end{aligned}$$

For q_2 , it again consists of two cases: 1) the second handoff is blocked and 2) the call is naturally terminated in the third cell; given that the mobile's first handoff is successful and that the call is not naturally terminated in the second cell. So q_2 equals

to q_1 multiplied by the probability that the first handoff is successful, i.e.,

$$\begin{aligned} q_2 &= q_1 \frac{\mu_R}{\mu_R + \mu_M}(1-p_f) \\ &= (1-p_f)\left(\frac{\mu_R}{\mu_R + \mu_M}\right)^2 \left[1 - \frac{\mu_R(1-p_f)}{\mu_R + \mu_M}\right]. \end{aligned}$$

Following the same derivation, we have

$$q_{i+1} = q_i(1-p_f)\frac{\mu_R}{\mu_R + \mu_M}$$

or

$$q_i = (1-p_f)^{i-1}\left(\frac{\mu_R}{\mu_R + \mu_M}\right)^i \left[1 - \frac{\mu_R(1-p_f)}{\mu_R + \mu_M}\right].$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments which helped to improve this paper.

REFERENCES

- [1] "Special issue: Introduction to mobile and wireless ATM," *IEEE Commun. Mag.*, Nov. 1997.
- [2] M. A. Marsan, C.-F. Chiasserini, R. L. Cigno, M. Munafo, and A. Fumagalli, "Local and global handovers for mobility management in wireless ATM networks," *IEEE Personal Commun.*, pp. 16-24, Oct. 1997.
- [3] K. Keeton *et al.*, "Providing connection oriented network services to mobile services," in *Proc. Usenix Symp. Mobile and Location Independent Computing*, 1993.
- [4] M. Karol, M. Veeraraghavan, and K. Y. Eng, "Implementation and analysis of handoff procedures in a wireless ATM LAN," in *Proc. GLOBECOM'96*, pp. 216-223.
- [5] S. J. Lee and D. K. Sung, "A new fast handoff management scheme in ATM-based wireless mobile networks," in *Proc. GLOBECOM'96*, pp. 1136-1140.
- [6] B. A. Akyol and D. C. Cox, "Rerouting for handoff in a wireless ATM network," *IEEE Personal Commun.*, pp. 26-33, Oct. 1996.
- [7] A. S. Acampora and M. Naghshineh, "An architecture and methodology for mobile-executed handoff in cellular ATM networks," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 1365-1375, Oct. 1994.
- [8] O. T. W. Yu and V. C. M. Leung, "Extending B-ISDN to support user terminal mobility over an ATM-based personal communications network," in *Proc. GLOBECOM'95*, pp. 2289-2293.
- [9] C. Rosado-Sosa and I. Rubin, "Jitter compensation scheduling schemes for the support of real-time communications," in *Proc. ICC'98*, pp. 885-890.
- [10] E. Chlebus, "Empirical validation of call holding time distribution in cellular communications systems," in *Proc. 15th Int. Teletraffic Congress*, June 1997, pp. 1179-1188.



analysis.

King-Sun Chan received the B.Eng. and M.Eng. degrees in electronic engineering from Shanghai University of Science and Technology (now Shanghai University), Shanghai, China, in 1989 and 1992, respectively, and the Ph.D. degree in information technology from the City University of Hong Kong, Kowloon, Hong Kong, in 1998.

He is now a Research Fellow in the Department of Electronic Engineering, City University of Hong Kong. His research interests are wireless communication systems, ATM switch design, and performance



Sammy Chan received the B.E. and M.Eng.Sc. degrees in electrical engineering from the University of Melbourne, Melbourne, Australia, in 1988 and 1990, respectively, and the Ph.D. degree in communication engineering from the Royal Melbourne Institute of Technology, Australia, in 1994.

From 1989 to 1994, he was with Telecom Australia Research Laboratories. Since 1994, he has been an Assistant Professor in the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong. His research interests are wireline and wireless ATM networks.



King-Tim Ko was born in Hong Kong. He received the B.Eng.(Hons.) degree in electrical engineering and the Ph.D. degree in communication engineering from the University of Adelaide, Australia, in 1977 and 1982, respectively.

From 1982 to 1986, he was a Senior Engineer with Telecom Australia Research Laboratories, Melbourne, Australia. In 1986, he joined the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong, where he is currently an Associate Professor. His main research interests include designs of communication networks as well as applications of multimedia.



Kwan L. Yeung (M'95–SM'99) received the B.Eng. and Ph.D. degrees in information engineering from the Chinese University of Hong Kong, Hong Kong, in 1992 and 1995, respectively.

In May 1995, he joined the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong, as an Assistant Professor. His main research interests include personal and mobile communication systems, high-speed networks, broad-band packet switches, and lightwave networks.



Eric W. M. Wong received the B.Sc. and M.Phil. degrees in electronic engineering from the Chinese University of Hong Kong, Hong Kong, in 1988 and 1990, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1994.

He joined the City University of Hong Kong, Kowloon, Hong Kong, as an Assistant Professor in 1994. His research interests are high-speed networks, video-on-demand, satellite communications, and dynamic routing. The most notable of these involved the analytical modeling of the least loaded routing scheme in circuit-switched networks. The model drastically reduces the computational complexity of designing and dimensioning telephone systems. The model has also served as a core for the analysis of many other advanced routing schemes, such as the real-time network routing recently implemented in the AT&T telephone network.