Various Handoff Schemes Used in the Handoff in ATM Networks

Amit Prakash      Ajay Kshemkalyani

University of Illinois at Chicago
Chicago- 60607

Distributed Computing Systems
Fall 2001
INDEX

1. Introduction

2. A brief overview of the ATM Network
   2.1 Handover Requirements in ATM Networks
   2.2 Mobility Functions

3. The Various Handoff Schemes Used for Handoff
   3.1 Nearest Common Node Rerouting
       Future Work and Recommendation
   3.2 Path Optimization for rerouting in the Inter Zone Handoff
       3.2.1 Optimal Path Computation
       3.2.2 Crossover Switch Determination
       3.2.3 Simulation Model for the Various Path Optimization Schemes
       Future Work and Recommendation

4. Discussion

5. References
Abstract

The world of telecommunication is rapidly changing. The capabilities of wireless networks are improving at a steady pace. This paper presents a design for implementing mobility into an Asynchronous Transfer Mode (ATM) network. The vision is ATM-based wireless telecommunication network that is capable of supporting a variety of today’s application with roam to grow for advanced application of the future. One of the major design issues in the Wireless ATM is the support of inter-switch handoff. This objective of this paper is to study the various rerouting schemes and the path optimization schemes required to perform the inter-switch handoff. We discuss here the various issues involved in the mobility in the ATM and then we study the various rerouting schemes that can be used for the handoff in an ATM Wireless Network.

1. Introduction

Wireless communication networks have rapidly grown to be a sizable part of the world telecommunications market over the past three decades. One reason for the growth of the Wireless Communication Networks is the mobility of a mobile terminal. The personal communication services provide mobility for the global use, a life-long user identification numbering system and the ability to accommodate varying data rates. An ATM network is designed to handle integrated voice, video, and data in a single protocol. The communicators aim to handle voice, data, and perhaps images. It is then logical to consider ATM network as the fixed supporting network since ATM can carry wireless traffic of varying data rates along long distances. [2,3] The Wireless ATM network has been designed to provide high-speed isochronous and asynchronous communication for wireless users, with the ever-growing demands of the users. The wireless companies aim to provide the users a hand held computer with conferencing capabilities and a data speed upto 10Mbits/sec [3]. This combined network is referred to as “ATM based Wireless PCS Network”. [1] We will concentrate here on implementing mobility within the fixed ATM network. Handoff occurs when a mobile terminal moves from one base station to another. ATM handoff differs from conventional voice handoff in that a mobile user may have several active connections with different bandwidth requirements and quality-of-service (QoS) constraints. The handoff function should ensure that all these ongoing connections are rerouted to another access point in a seamless manner. In other words, the design goal is to prevent service disruption and degradation during and after the handoff process. Handoff is implemented by the network to give the users freedom of motion beyond a limited wireless coverage area while they are communicating. The handoff is the procedure by which a user’s radio link is transferred from one radio port to another through the network without an interruption of the user connection [4]. Handoff procedure is performed to assure the integrity of a radio connection and to minimize interference to the users in the coverage area of neighboring cells [4,5].

In this paper we discuss a briefly about the ATM Networks in the next section. Later we discuss about the rerouting and the path optimization schemes required for introducing the concept of mobility in the ATM Networks.
2. A Brief Overview of the ATM Network

The wireless ATM network consists of radio ports, user terminals and network interface equipment. A user terminal might have a few simultaneous connections in the Wireless ATM network. When a handoff occurs these connections may need to be rerouted. A group of radio ports is connected to the same Wireless ATM network interface equipment. This collection of ports is called a “ZONE”[1]. The zone architecture is illustrated in the figure 1 below. The “Zone Manager” process manages the zone. There are two levels in the handoff event:
   a) Network Level
   b) Radio Level

The radio level handoff is the actual transfer of the radio link between two ports. The Network level supports the radio level handoff by performing the rerouting and buffering. The radio level handoff determines some processes used in the network level handoff. It is assumed that these zones are interconnected by wireless ATM network switching nodes. Based on the “ZONE” concept the Handoff can is classified as:
   a) Intra Zone Handoff
   b) Inter Zone Handoff

As the name suggests in the intra zone handoff the user is moving within the zone. The only rerouting that is performed in this case is the Wireless ATM network interface equipment within the zone. This type of rerouting doesn’t involve ATM network switching.

While the Inter-zone handoff occurs when the radio ports involved in the handoff belong to different zones. In this case the wireless ATM network is involved in the rerouting. The number of switched involved in the rerouting depends on the location of the handoff and the topology.

Figure 1. A Zone in the Wireless ATM Network
2.1 Handover Requirements in ATM Networks

The requirements for the handover procedure are listed in ATM Forum/96-989. The following sections expand and detail further the original handover requirements list.

2.1.1 Handover Latency
• The total time for the completion of handover should be appropriate for the rate of mobility of the WT. The handover process should be fast enough so that the handover decision is still valid for the new WT position after the handover process is complete.
• The switching of the active VCs from the old data path to new data path should be as efficient as possible in order to minimize the interruption to cell transport. This can be a considerably shorter time than that for the overall handover process.

2.1.2 Scalability
• The handover procedure should support seamless handover between APs in a private network.
• It should support handover between APs in different private networks connected by a public network. (Billing may be an issue across the public network.)
• It should support handover between APs in public networks.
• The handover procedure should also work on networks that contain non-mobile-enabled switches.

2.1.3 Quality of Service (QoS)
• The handover procedure should aim to preserve the requested QoS of all VCs at handover. This may not always be possible and some form of QoS re-negotiation and/or dropping of certain VCs on a priority basis may be required.
• Information about resource allocation at a candidate AP should be advertised to the WT and to other APs in the network. A handover algorithm can then use this information to select the AP that will best support the requested QoS for active VCs. This information will also be useful for resource load balancing between radio adjacent APs.

2.1.4 Signaling Traffic
• Handover-signaling traffic should be kept to a minimum in order to reduce the load on the wired network and the air interface to the WT.
• The type of signaling protocol, whether user plane or control plane, used for handover should take into consideration the processing load on intermediate switches.

2.1.5 Buffer Tradeoff
• The handover procedure should not require modifications to existing network switch buffering hardware implementations.
• There is no strong requirement for minimal buffering in itself; rather there should be a trade off between buffer delay and cell loss or cell reordering at handover depending of the traffic class of the active VC.

2.1.6 Data Integrity
Minimize cell loss but more importantly avoid cell duplication or cell re-ordering.

2.1.7 *Group Handover*

- The handover procedure should support the efficient handover of multiple active VCs. This will be affected by the handover approach used.

### 2.2 Mobility Functions

#### 2.2.1 Location Management

Connection setup protocols in ATM (UNI/NNI signaling) have implicitly assumed that endpoints of a connection refer to static terminals, i.e. it is not necessary to determine dynamically a terminal’s current attachment point to the network before attempting to establish a connection to that terminal. However, with mobile ATM terminals, the location of such a terminal with respect to the network may no longer be deduced from its endpoint address. Additional addressing schemes and protocols are needed to locate and track mobile terminals, along with suitable modifications to the connection setup process.

![Figure 2. Mobile ATM Network Concept](image)

For example, in figure 2, a connection is setup to the mobile in each of the two cells (under radio ports b3 and b4). Under a possible addressing architecture, the name m is resolved to an address b3.x, that is specific to radio port b3, in order to setup the first connection. The mobile then moves to the cell under radio port b4. When the second
It is not necessary for a caller to determine a priori whether
(a) The endpoint for connection setup is mobile and
(b) The mobile is presently attached to its home switch or a foreign switch.

The caller initiates a connection setup to the mobile using its home address. The SETUP message is routed to the mobile’s home switch, as before. But now, if the mobile is attached to the network via a switch other than its home, its home switch returns a response to the caller indicating the mobile’s current foreign address. This could be sent as a "cause" for failure to setup the connection using the mobile’s call reference: consequently, the resources reserved for the connection (during the forward path) are released on the return path of the response. As a result of this response, the caller receives the mobile’s current foreign address. The caller can then initiate a connection setup procedure using the mobile’s foreign address, which is then routed to the mobile’s current network attachment point.

2.2.2 Handoff Control

Once a connection has been established between a pair of endpoints, current protocols assume that the connection path does not change during the period of a connection lifetime (except due to failures of switches and links). This assumption is invalidated when the endpoint is mobile. The path of the connection for a mobile user needs to be frequently reset up as the mobile user moves from one cell to another. For example, in figure 2, the connection that was setup to the mobile terminal through b3 is handed-off to b4 by deleting the sub-path from switch S3 to b3, and augmenting the truncated path from S3 to b4. An efficient handoff control protocol does dynamic path re-routing instead of whole path re-establishment. Handoff schemes can be classified into path re-routing and path extension. First scheme is based on removing a part of the existing connection and adding a new sub-path from the point of detachment (the "crossover" point). It is necessary to determine a crossover point in this scheme. How to select the crossover point is a fundamental problem in handoff control techniques and different methods give different performance for handoff control in terms of latency, data loss and resource utilization. The second scheme is based on extending the original connection from previous access point to new access point. This scheme is currently being used in the telephony networks. Although it looks relatively simple for a single handoff, the overall process has to include other complicated algorithms for loop removal and route optimization.

3. The Various Schemes Used for Handoff

3.1 Nearest Common Node Rerouting (NCNR)

We study here one of the rerouting schemes used in mobile ATM networks-- Nearest Common Node Rerouting (NCNR). This attempts to perform the re-routing for a handoff
at the closest ATM network node that is common to both zones involved in the handoff transaction. The term “common” is used to denote a network node that is hierarchically above both of the zones in question or a parent of both zones in the network topology tree [6]. This scheme minimizes the resources required for the rerouting and conserves the network bandwidth by eliminating unnecessary connections. As mentioned in the introduction, the users of the Wireless ATM network may subscribe to services ranging from time sensitive traffic types (audio, video) to throughput dependent traffic types (data, file transfer, etc). The zone managers know the traffic type of the connection involved in the handoff. The two kinds of traffic impose differing constraints on the network and the handoff process. For example the time-sensitive (TS) voice traffic will not be easy to buffer due to constant cell generation rate and strict time delay constraints; however, it can tolerate occasional loss of cells. On the other hand, data traffic will not tolerate cell loss, but may tolerate delays on the order of few milliseconds. Based on this fact, the NCNR procedure [6] for these two types of traffic is different. Due to the nature of the fixed network, the transmission delay and latency of the links from the NCN to the zones involved in the handoff are assumed to be negligible compared to the radio transmission medium. The cell sequence in the NCNR scheme is preserved. For the TS traffic, the cell sequence in NCNR is preserved by discarding the cells, duplicated in the NCN and transmitted to both zones involved in the handoff, at the zone that is not currently in contact with the Handoff Terminal. This reduces the number of active paths to one in the downlink, thereby preserving the cell sequence. In the uplink, data from only one zone are forwarded to the endpoint by the NCN. This preserves the cell sequence.

For the throughput dependent (TD) traffic, during a handoff, the involved zone managers start buffering information from the start of the radio level handoff to the completion (successful or unsuccessful); moreover, until the radio handoff is started, the user connection is assumed to be active and all the incoming cells are transmitted to Handoff terminal (HT) through the previous port. If that is not possible, the cells are forwarded to the candidate port through the NSN for later transmission. In the reverse direction, the HT buffers the user information as soon as the radio handoff is initiated. The cells that are in transmission in the network from the HT to the end point are delivered normally.

By employing the procedure discussed above for TD traffic, the cell sequence during the rerouting is preserved due to the fact that at any given time during a handoff, there is only one active path between the two parties involved in the connection. As there is only one active path at any given time, all the cells transmitted through the network take this path and arrive in sequence to their destination.
NCNR for TS Traffic

The NCNR for the TS traffic is performed as follows:

1. A handoff session between zones A and B is started. Let B be the candidate for the handoff. Let A be the present zone.

2. The zone manager of A first checks to see if a direct physical link between A and B exists. There are two possible cases if this condition is satisfied:

   a) If A is the parent of B, then A notifies B and the new connection is established without any further network involvement. After the connection is established, A acts as an anchor for the connection. Until the stability of the handoff is established, both A and B act as a network connection points for the user connection. This process is explained at the step 6 of this procedure. Once the radio level handoff is completed, and then A acts only as a wireless ATM switch in the connection path.
b) If B is the parent of A, then A sends a message to B relaying the handoff request. B then acts as an anchor for the handoff procedure. Until the stability of the handoff is established, both A and B can be used for information transfer from/to the terminal to/from the network. Once the handoff is stable, B deletes the user connection from itself to A. The rerouting is thus completed.

3. If A and B are not connected by a physical link then the zone manager of A (ZMA) contacts the end point for the user connection by sending a handoff start message. The handoff start message contains the ATM addresses of the zones A and B and the endpoint for the user connection.

4. The handoff start message traverses the network from A to the endpoint for the user connection. The network switching nodes on this path upon receiving this message check to see whether all three addresses are routed on different egress ports of the switch. When such a node is found it is designated as a NCN. The NCN sets the NCN bit in the handoff start message. The rest of the switches on this path do not perform the egress port test.
5. The NCN then forwards the reroute message to all of the switches located between B and itself. The nodes that receive the reroute message check for resource availability; if the resources are available then the connection is established and circuit translation table is set up. Else if the resources are not available then, the handoff attempt fails and the involved parties are notified.

6. When B receives the reroute message, a reroute acknowledgement message is sent from B to A. This message completes the rerouting process. The radio level handoff is attempted at this point. As the radio level handoff is started, the NCN starts to forward the user information to both A and B in a point to multipoint manner. This multi party connection is necessary until the radio level handoff is stabilized.
7. If the radio level handoff is successful, the connection between A and NCN is cleared by A by sending a clear connection message to the NCN. The data buffered at B is then transmitted in sequence to the user terminal. If the radio level handoff is not successful, then the information buffered at A is transmitted to the user terminal.

*Recommendation and Future Work:*

The above scheme doesn’t consider the problems of bandwidth consumption due to intermediate routes and the QoS maintenance due to the presence of multiple connections of different traffic types being served by a single MT. In the future, we should give the source switch the responsibility of determining the new route to the moving endpoint, regardless of which endpoint has moved.

The above-mentioned protocol also doesn’t study about the delay sensitive connections and loss sensitive connections. In future studies, for delay sensitive traffic, we should aim to achieve a fast handoff route. Thus, the endpoint switch uses connection extension to reach the MT’s next hop location. An optimal route would be then calculated and the delay sensitive connections would be switched from the extended route to the optimal route. For loss sensitive traffic, we should aim to preserve packets. Thus, when the MT moves to a new location, the packets on the loss sensitive connection would be buffered and then later switched to the optimal route.

### 3.2 Path Optimization for rerouting in Inter Zone Handoff

As mentioned before, for inter-switch handoff, connection rerouting is also required within the fixed ATM network. Several connection-rerouting protocols to facilitate inter-switch handoff have recently been proposed [9,10]. Two methods based upon a partial connection re-establishment are the *path extension* and the *path rerouting* schemes. The rationale behind path extension is to extend the original connection to the switch to which the new base station is connected. Referring to figure 8, the switches to which the original and new base stations are connected are usually referred to, respectively, as the *anchor switch* and the *target switch* [8]. The path extension method extends the connection from the anchor switch to the target switch during handoff. The minimum hop path between these two switches is usually chosen as the extended path. The path extension scheme is fast and simple to implement. QoS degradations such as cell loss, duplicate cells, and missequence cells do not occur. However, since the extended path is longer than the original one, certain QoS requirements, such as cell transfer delay and cell delay variation, may not be guaranteed after a handoff. In addition, data looping may occur when the mobile terminal moves back to the previous anchor switch later, which leads to an inefficient use of network resources.

Path rerouting can be considered as a generalization of the path extension scheme. In path extension, the anchor switch extends the original connection to the target switch, while in path rerouting; any switch along the original connection can be selected to set up a branch connection to the target switch (see figure 9). The switch chosen to perform this
function is usually referred to as the *crossover switch* [9] or the *handoff switch* [11]. Depending on the performance criteria of the crossover switch discovery algorithms, the end-to-end path after rerouting may not be optimal. The two-phase handoff protocol, which combines the advantages of path extension and path rerouting schemes consists of two stages, namely: path extension and possible path optimization.[10] Referring to figure 10, path extension is performed for each inter-switch handoff, path optimization is performed once a while. During path optimization, the network determines the optimal path between the source and the destination (i.e., the path between the remote terminal and the current anchor switch in figure 6), and transfers the user information from the old path to the new path. The major steps during path optimization execution generally involve:

1. Determining the location of the crossover switch;
2. Setting up a new branch connection;
3. Transferring the user information from the old branch connection to the new one;
4. Terminating the old branch connection.

![Figure 8. Path Extension Scheme](image)

![Figure 9. Path Rerouting Scheme](image)
Since the mobile terminal is still communicating over the extended path via the current base station while path optimization takes place, this gives enough time for the network to perform the necessary functions while minimizing any service disruptions. Notice that the path optimization process described above is not restricted to the two-phase handoff protocol, but it can also be applied to other connection rerouting protocols where the location of the crossover switch is not the optimal one. In addition, when the mobile terminal moves to another switch during the execution of path optimization, path extension can still be used to extend the connection to the target switch. Although path optimization can increase the network utilization by rerouting the connection to a more efficient route, transient QoS degradations such as ATM cell loss and an increase in cell delay variation may occur. In addition, path optimization increases the processing load of certain switches and increases the signaling loads of the network. Thus, path optimization after each path extension may not be necessary or desirable. To ensure a seamless path optimization, two important issues need to be addressed:

### 3.2.1 Optimal Path Computation

The handoff reference model described for the mobile-to-mobile connection, the fixed segment remains unchanged during its connection lifetime. Hence, instead of determining a new end-to-end route between the source and the destination, only the computation of the affected handoff segment is necessary. To facilitate the computation of the handoff segment, the cell transfer delay constraint in the wired ATM network is partitioned into three delay bounds, namely, the delay constraint in the fixed segment, and the delay constraints in the two handoff segments (between the boundary switch and the anchor switch). Thus, for a mobile-to-mobile connection, the path optimization procedures of the two-handoff segments are independent to each other. During the optimal path computation, the path between the current edge switch and the boundary switch is determined based on the BDC path algorithm [13].

### 3.3.2 Crossover Switch Determination

After the new handoff segment is computed, the anchor switch determines the location of the crossover switch in a centralized manner. Based on the information from the routing information database, the anchor switch compares the original and the new handoff.
segments to determine the location of the crossover switch. Starting from the boundary switch, the crossover switch is the one where the original and the new handoff segments begin to diverge. As an example, in Figure 11, where the connection in the upper portion represents the original handoff segment, and the connection in the lower portion represents the new handoff segment, switch SW 2 is the location of the crossover switch. The original and the new handoff segments can be the same if the boundary switch happens to be the crossover switch. In this case, the current handoff segment is the optimal path. Thus, no further path optimization executions are necessary.

![Figure 11. Crossover switch determination.](image)

1. How to minimize service disruptions during path optimization?
2. When and how often should path optimization be performed?

We discuss the first issue using a signaling protocol in which buffering is used at the anchor switch and the crossover switch during path optimization to prevent ATM cells loss and maintain cell sequencing. The second issue is discussed later and we talk about three path optimization schemes, namely: exponential, periodic, and Bernoulli. These schemes are not optimal but they are simple to implement. An analytical model and a simulation model were used to analyze the performance of these path optimization schemes. For the discrete time analytical model, closed-form solution of the expected cost of a mobile call and the optimal operating point are derived for each path optimization scheme. Simulations based on random graphs corroborate the analytical results. Path optimization schemes are grouped into four types, namely,

(i) QoS-based,
(ii) Network-based,
(iii) Time-based, and
(iv) Handoff-based.

**QoS-based schemes**

As the name implies, QoS-based path optimization schemes trigger path optimization of each mobile connection based on its current QoS measures. For example, path optimization can be initiated if the number of hops of the path is greater than a certain number, or if the end-to-end cell transfers delay bound is violated. To implement those QoS-based path optimization schemes, information about the quality of the current path in terms of the defined QoS measures (e.g., hop count, current average delay) must be maintained by the network.
Network-based schemes
Network-based path optimization schemes trigger path optimizations for a group of connections based on the existing traffic load of a switch or the utilization of the network. For example, a network switch can initiate path optimization for a group of mobile connections whenever the new call dropping probability of a certain traffic class exceeds a particular threshold. During path optimization, a number of connections will be rerouted to some other switches, thereby reducing the traffic load of that switch.

Time-based schemes
In this case, path optimizations are triggered at time instants, which are independent of the current QoS of the connection or the network utilization of the network. The time instants can be deterministic or random. For example, the time between path optimization can be based on some random processes. In addition, it can also be a function of the velocity of the mobile terminal, the dwell time, and the residual service time of the mobile connection. Two of these time-based schemes, which are simple to implement:

1. **Periodic path optimization scheme.** The time to perform path optimization is periodic with period T.
2. **Exponential path optimization scheme.** The time between path optimizations is modeled as an exponentially distributed random variable with mean $1/\nu$.

The periodic scheme has been proposed within the ATM Forum wireless ATM working group [15] to facilitate the inter-switch handoff protocol. Notice that both periodic and the exponential schemes trigger path optimization independent of whether or not there is an inter-switch handoff. Thus, unnecessary path optimizations may be performed for stationary mobile connections.

Handoff-based schemes
The handoff-based path optimization schemes trigger path optimization for each mobile connection based on some criteria after each inter-switch handoff. For example, it can be a function of the number of previous hand-offs, the velocity of the mobile terminal, and the residual service time. The Bernoulli Path Optimization Scheme is also analyzed below:

3. **Bernoulli path optimization scheme.** After each path extension, there is a probability $p$, $0 <= p <= 1$, such that path optimization is performed. For the Bernoulli scheme, path optimization may only occur on condition that there is an inter-switch handoff. Thus, path optimization is never triggered after each hand-off if $p=0$. On the other hand, path optimization is always performed after each path extension if $p=1$. Path optimization can also be invoked based on a combination of the above schemes.

3.3.3 Simulation Model for the Various Path Optimization Schemes
In this section, we present the framework of the simulation model for the comparisons between different path optimization initiation schemes. An ATM network is modeled as a
random graph based on the minimum weight-spanning tree. The generation of the random graph consists of the following steps [17]:

1. $V$ nodes are randomly distributed over a rectangular coordinate grid. Each node is placed at a location with integer coordinates. A minimum distance is specified so that a node is rejected if it is too close to another node. The Euclidean metric is then used to calculate the distance $d(i,j)$ between each pair of nodes $(i,j)$.
2. A fully connected graph is constructed with the link weight equals to the Euclidean distance.
3. Based on the fully connected graph, a minimum weight-spanning tree is constructed.
4. To achieve a specified average node degree of the graph, edges are added one at a time with increasing distance.

If node $i$ and $j$ are connected, then the link delay, denoted as $d_{ij}$, is assumed to be equal to

$$d_{ij} = d(i,j) + \delta$$

(1)

Where $\delta$ is a uniformly distributed random variable in the range $0 \leq \delta \leq \delta_{\text{max}}$. In equation (1), the first term can be interpreted as the propagation delay of the link, and the second term approximately models the queuing delay of the link.

![Figure 12. A 20-node random graph with node degree of 3; minimum distance between any two nodes is 15](image)

As an example, a 20-node random graph generated from the above model is shown in figure 12. The size of the rectangular coordinate grid is 100X100. The minimum distance between any two nodes is 15. The average node degree of the graph is 3. For equation (1), $\delta_{\text{max}}$ is set to be equal to 100. Each node represents an ATM switch and each edge represents a physical link connecting the two switches.

For the simulations, ten different 20-node random graphs, similar to the example shown in figure 7, can be generated. Each random graph is employed in 10000 simulation runs. For each simulation run, a call is generated with two nodes chosen randomly as the
source and destination nodes. Dijkstra’s algorithm is used to compute the shortest delay path between these two nodes. The source node is assumed to be stationary. The destination node becomes the anchor switch of the mobile connection. The call duration and the time between inter-switch handoff are modeled as exponentially distributed random variables with mean $1/\mu$ and $1/\lambda$, respectively.

During each inter-switch handoff, the target switch is restricted to be one of the neighboring switches of the current anchor switch. Path extension is used to extend the connection from the anchor switch to the target switch. Subsequent path optimizations may be triggered based on different initiation schemes as described previously. Similar to the performance metric used in our analytical model, the performance metric in our simulation model is the average cost per call, which is defined as the sum of the link cost and the signaling cost due to inter-switch handoff, and is given by

\[
\text{Cost} = \eta_{PE} C_{PE} + \eta_{PO} C_{PO} + \sum_{i=1}^{m} \tau_i l_i C_{\text{link}}
\]  

(2)

Where $\eta_{PE}$ and $\eta_{PO}$ denote the number of path extension and path optimization, respectively, during a call. $C_{\text{link}}$, $C_{PE}$ and $C_{PO}$ are as defined in the previous section; $m$ is the number of events per call; $\tau_i$ denotes the time elapsed between event $i-1$ and $i$; and $l_i$ denotes the number of links between two end nodes during $\tau_i$. The average cost per call is calculated by averaging the total cost per call over all $(200000)$ simulation runs.

The simulation model described above can be extended to handle other network topologies (e.g., star, ring, and hierarchical), as well as other statistical distributions for the call duration and the time between inter-switch handoff.

**Recommendation and Future Work:**

We need to know when do we need to perform the path optimization for the two-phase handoff protocol. In future we would study that path optimization needs to be performed when the number of links of the path is greater than a certain threshold. Future work on this topic can be working on the QoS constraints during the path optimization. We also need to think about the multi-cast connection in which a group of mobile users are communication with each other.

**4. Discussion**

This paper presents the various schemes that can be used for handoff in the ATM networks. We have presented the Nearest Common Node Rerouting technique. This attempts to perform the rerouting for a handoff at the closest ATM network node that is common to both zones involved in the handoff transaction. Path optimization may be necessary if the rerouted path is not optimal. In this paper, we have considered the scenario where the path extension is used for each handoff, and path optimization is invoked whenever the delay of one of the potential handoff paths exceeds the delay constraint. The path optimization signaling protocol ensures that the uplink and downlink ATM cell sequencing is maintained by buffering at the anchor switch and the crossover switch. Then, we present the various path optimization schemes used for the rerouting during the handoff. In the situation that the handoff domain consists of a number of peer groups, the design of a centralized or distributed crossover switch algorithm is a future issue.
5. References


