Handoff Rerouting Scheme for Multimedia Connections in ATM-based Mobile Networks

Janise McNair   Ian F. Akyildiz
Broadband & Wireless Networking Laboratory
School of Electrical & Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
{mcnair, ian}@ee.gatech.edu

Sridhar Radhakrishnan   Gopal Racherla
School of Computer Science
University of Oklahoma
Norman, OK 73019 USA
{gopalr,sridhar}@cs.ou.edu

Abstract

Future wireless networks promise multimedia communication and computing services for both fixed and mobile users. One of the most important challenges of wireless multimedia service for mobile users is maintaining a guaranteed quality of service over end-to-end connections with multiple mobile endpoints. In this paper, we describe a new handoff rerouting scheme that supports multimedia traffic in a connection-oriented mobile network, such as Mobile ATM. We analyze and compare the performance of the new scheme with respect to three related handoff rerouting schemes for connection-oriented wireless networks. The new technique shows a superior performance in route efficiency and bandwidth consumption for all traffic types, and a smaller handoff delay for delay sensitive traffic.

1. Introduction

Future personal communications services (PCS) networks promise personalized, multimedia-based communication and computing services to both fixed and mobile users. One of the important challenges of PCS is to maintain a guaranteed level of quality of service (QoS) for mobile and wireless connections. While wireline networks have the advantage of plentiful bandwidth and reliable connections between fixed endpoints, future wireless systems must compensate for limited bandwidth, error-prone wireless links, and connections where multiple endpoints are roaming. The process of transferring the active connections of a roaming mobile user to a new channel or cell is called handoff, or handover [1]. Handoff rerouting attempts to achieve optimal routes for the handoff connections, while minimizing bandwidth consumption, packet loss and handoff delay. Mobile-mobile connection rerouting increases the challenge of supporting multimedia traffic, since it requires additional intermediate routes that increase bandwidth use, cause additional handoff delays, and introduce greater opportunities for packet loss.

Research on handoff re-routing for connection-oriented wireless networks has evolved from two extreme goals. The first extreme is to ensure an optimal route by establishing an entirely new connection route each time the mobile terminal (MT) moves to a new location. Although an optimal route is obtained, the handoff latency is greatly increased due to the connection establishment process. The other extreme is to ensure the smallest handoff delay by keeping the original connection route, and then extending the route hop-by-hop as the MT moves. This method provides the fastest extension, but is a wasteful use of network resources for many hops. Between these two extremes, three types of connection-oriented handoff re-routing schemes are described in recent research: (1) *anchored (or static/dynamic) handoff routes*, (2) *multicast handoff routes*, or (3) *chained handoff routes*.

In [8], an anchored handoff route is used. In this technique, an anchor switch is designated for each MT in the original connection route. During the lifetime of the connection, the anchor segment (the path that connects the anchor switches) never changes. Then, when one of the MTs moves to a new location, only the portion of the connection from the MT to its anchor switch needs to be rerouted. This method prevents the need to establish an entirely new route for each handoff. However, the anchor segment may cause longer routes or looped routes if static, or higher complexities and packet loss if dynamic. In [4], a multicast handoff rerouting technique is discussed. Incoming traffic for the MT is multicast to several surrounding base stations, which buffer the packets until the MT enters the cell. This scheme greatly reduces handoff delay, but expends multicast bandwidth for each user, and requires heavy buffering in the cells where the user is not located. In [5, 7], a chained handoff rerouting technique is explored. When the MT moves, the connection is first extended to the next cell until an optimal

0-7803-6507-0/00/$10.00 ©2000 IEEE

route can be calculated. Then the extended route is switched to the optimal route. This scheme requires several intermediate routes that consume bandwidth for each movement of the MT. None of the above schemes distinguish multimedia traffic requirements, and they give little treatment to future issues, such as MTs with multiple active connections or connections with multiple mobile endpoints.

In this paper, we describe a new handoff rerouting scheme that supports multimedia traffic over mobile connections in a connection-oriented wireless network, such as Mobile ATM. In Section 2, we introduce the new handoff rerouting technique. Next, in Section 3, we analyze and compare the performance of the new scheme with respect to the related work. Section 4 presents the numerical results, followed by the conclusion in Section 5.

2. New Handoff Rerouting Scheme

The new handoff rerouting scheme serves multimedia traffic by grouping connections into two categories: delay sensitive connections and loss sensitive connections. The switches at the endpoints of the connections keep track of each traffic category. When one of the endpoints of a mobile connection moves, the endpoint switches are able to reroute the connection according to the nature of the traffic. For delay sensitive connections, the goal is to achieve a fast handoff route. Thus, the endpoint switch uses connection extension to reach the MT's next hop location. An optimal route is then calculated, and the delay sensitive connections are switched from the extended route to the optimal route. For loss sensitive traffic, the goal is to preserve packets. Thus, when the MT moves to a new location, the packets on the loss sensitive connections are first buffered and then later switched to the optimal route.

2.1. Procedure Description

To describe the connection rerouting procedure, we consider an original connection path from a source MT, MT_S, to a destination MT, MT_D. MT_S resides at the original source switch, SW_S, while MT_D resides at the original destination switch, SW_D. The handoffs occur between the network base stations. However, since our scheme focuses on the routing activity that occurs at the network switches, we describe the effect of the MT moving between network switches. In [8], it was suggested to provide base stations with limited switching capabilities. In that case, we would replace the switches in our illustration with the enhanced base stations.

First we describe the rerouting procedure at the source switch, highlighting the source's responsibilities when MT_S moves. Then we describe the procedure at the destination switch to highlight the responsibilities for the destination switch when MT_D moves.

2.2. Procedure at the Source End of the Connection

The responsibility of the source switch is to establish the handoff connection when either MT_S or MT_D moves to a new location. The procedure from the perspective of the source switch is shown in Figure 1(a)–(d) and outlined below.

1. Figure 1(a). MT_S moves in to the service area of $SW_{S,NEW}$. Then MT_S notifies its former switch, $SW_{S,OLD}$, that it will handoff to the new switch.
2. Figure 1(b). $SW_{S,OLD}$ sends the list of MT_S's active connections to $SW_{S,NEW}$. $SW_{S,OLD}$ buffers the loss sensitive traffic and extends the delay sensitive connections to $SW_{S,NEW}$. ($SW_{S,NEW}$ buffers MT_S's upstream loss sensitive traffic.)
3. Figure 1(b). $SW_{S,NEW}$ attempts to reroute the handoff connection to the original destination switch, SW_D. If MT_D is still located at SW_D, then SW_D accepts the new path request and the process continues from Step 6).
4. Figure 1(c). If MT_D is no longer located at SW_D, SW_D rejects the path request and sends a redirect message back to $SW_{S,NEW}$ that MT_D is now located at a new destination switch, $SW_{D,NEW}$.
5. Figure 1(c). $SW_{S,NEW}$ attempts to reroute the handoff connection to the new destination switch, $SW_{D,NEW}$.
6. Figure 1(d). $SW_{D,NEW}$ accepts the path request.
7. Figure 1(d). $SW_{S,NEW}$ notifies $SW_{S,OLD}$ to transmit the buffered traffic. Then $SW_{S,NEW}$ begins to send all of the traffic connections on the new path, and tears down the extended path through $SW_{S,OLD}$.

As mentioned previously, the source switch is responsible for rerouting the handoff connection when MT_S or MT_D moves. The responsibility of the destination switch is to inform the source switch of changes in the location of MT_D and to redirect handoff connection requests to the new destination switch, if necessary. We now describe the procedure for the case when the destination MT, MT_D, is the first endpoint to move to a new location.

2.3. Procedure at the Destination End of the Connection

The handoff rerouting process at the destination is shown in Figure 2(a)–(d) and explained below.

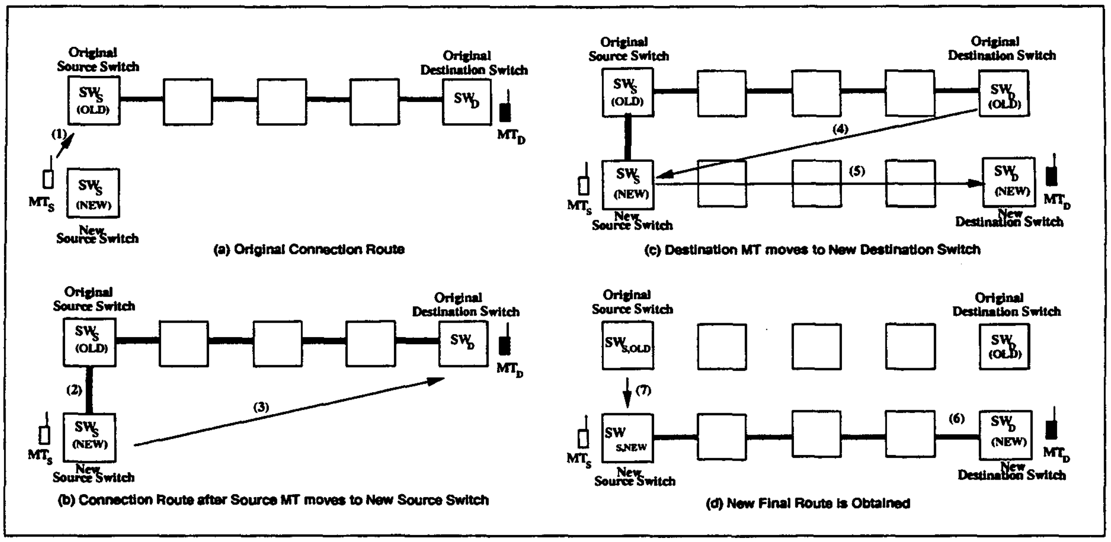

Figure 1. Handoff Rerouting Procedure at the Source

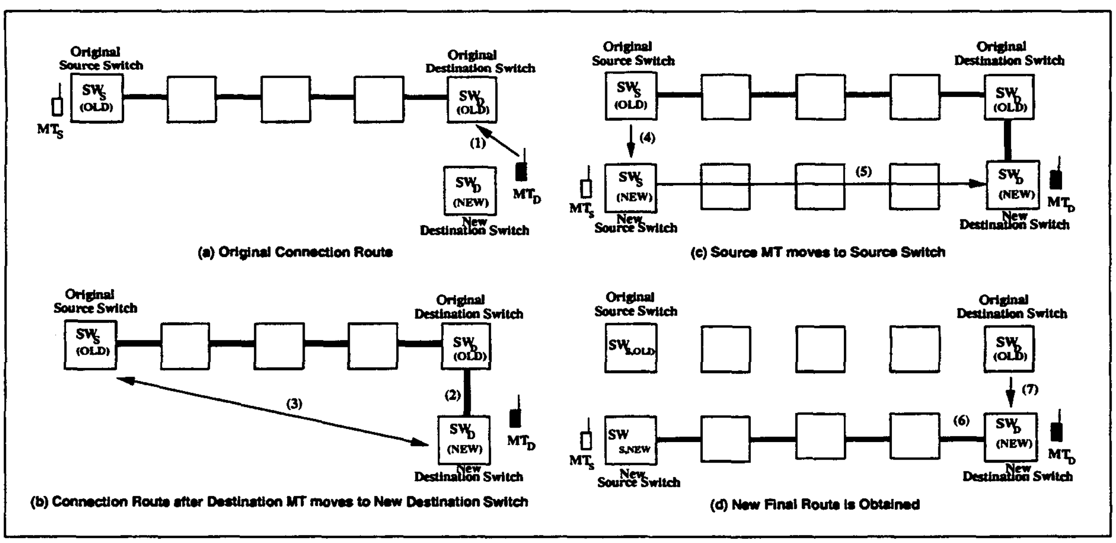

Figure 2. Handoff Rerouting Procedure at the Destination

1. Figure 2(a). MT_D moves into the service area of $SW_{D,NEW}$. Then MT_D notifies its former switch, $SW_{D,OLD}$, that handoff has been initiated to $SW_{D,NEW}$.
2. Figure 2(b). As in the source switch procedure, $SW_{D,OLD}$ extends the delay sensitive connections to $SW_{D,NEW}$ and buffers the loss sensitive connections. ($SW_{D,NEW}$ begins to buffer MT_D's upstream loss sensitive traffic.)
3. Figure 2(b). $SW_{D,OLD}$ then notifies the original source switch, SW_S, that MT_D has moved to $SW_{D,NEW}$. If MT_S is still located at SW_S, SW_S attempts to reroute the handoff connection to $SW_{D,NEW}$ and the process continues from Step 6.
4. Figure 2(c). If MT_S has already moved to a new location ($SW_{S,NEW}$), SW_S sends a message to $SW_{S,NEW}$ to redirect all of MT_S's connections to $SW_{D,NEW}$.
5. Figure 2(c). $SW_{S,NEW}$ then attempts to reroute the handoff connection to $SW_{D,NEW}$.
6. Figure 2(d). $SW_{D,NEW}$ accepts the path request.
7. Figure 2(d). $SW_{D,NEW}$ notifies $SW_{D,OLD}$ to transmit the buffered traffic. $SW_{D,NEW}$ begins to send all of the traffic connections on the new path, and tears down the extended path through $SW_{D,OLD}$.

The new scheme results in an optimal path from the source switch to the destination switch and prevents the establishment of intermediate handoff routes. In addition, the delay sensitive traffic is rerouted quickly, while the loss sensitive traffic is preserved in buffers. To join the separate traffic streams into the optimal path, the buffered loss sensitive traffic must be transmitted from the buffer and delay sensitive traffic must be switched from the extended route to the optimal route. Combining several sources into a single route may cause a jitter effect known as *negative jitter*. Although the negative jitter problem is out of the scope of this paper, the interested reader can find an investigation of the problem in [3].

As discussed in the introduction, several related techniques have been proposed to address the handoff rerouting problem. In the next section, we evaluate the performance of the new scheme with respect to related schemes.

3. Handoff Rerouting Protocol Comparison

3.1. Framework

To compare the performance, we consider the case of an original connection between a mobile source and a mobile destination, as shown in Figure 3. The solid line at the top of Figure 3 represents the original connection. The dotted lines indicate the routing distance between each switch. For simplicity, we normalize the horizontal and vertical switch-to-switch distance to one hop. MT_S and MT_D move simultaneously for m movements until each MT reaches its final switch. The movements are consecutive in one direction, and the handoff rerouting does not complete until the mth movement. Each move occurs between two separate switching elements, and requires the network to continue rerouting until the final route is established.

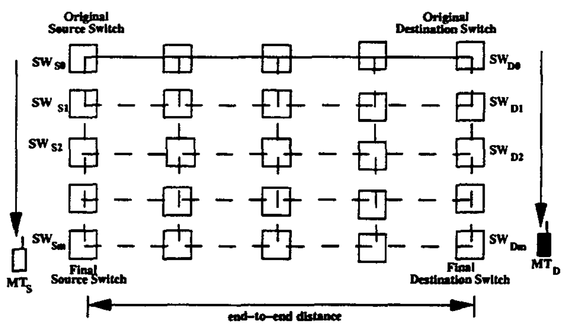

Figure 3. Interconnected Network of Switches

We compare the new technique to the static/dynamic anchored segments technique [9, 2, 8], the multicast group technique [4] and the chaining technique [5, 7], according to the final route length, the summed route length, the handoff delay, and the bandwidth consumption.

3.2. Final Route Length

Each of the handoff rerouting protocols is followed for m movements of the source and destination MT. The final route length, R_f, is the number of hops from the final source switch to the final destination switch:

$$R_f = \sum_{i=1}^{N} d_i, \qquad (1)$$

where N is the total number of switches in the final connection path, and d_i is the distance between switch $(i-1)$ and switch (i). For an optimal route, R_f is equal to the end-to-end distance (Figure 3).

3.3. Summed Length of the Intermediate Routes

We now consider the resources used by intermediate routes for the rerouting schemes that establish routes each tear the route down if one of the MTs moves again before communication can begin. The summed route length, R_s, is a measure of the resources that are used for repeating the set up/tear down process:

$$R_s = \sum_{j=1}^{m} \tilde{r}_j, \tag{2}$$

where m is the total number of movements made by the source and destination MT, and $\tilde{r}_j$ is the intermediate route length established as a result of move j. Note that $\tilde{r}_j$ in (2) includes only that portion of the route that changes as the MTs move. For example, since the static-dynamic anchored handoff rerouting technique keeps the same anchor segment within the connection route for the duration of the call, the anchored segment is not an additional resource that must be assigned. Thus, it is not included in $\tilde{r}_j$. However, the dynamic portions are included in $\tilde{r}_j$, since they must be re-assigned.

3.4. Bandwidth Consumption

Once the bandwidth consumption due to intermediate routes is obtained, the overall bandwidth consumption, BW, can then be found by multiplying the number of intermediate routes for each protocol by the wireless and wireline bandwidth per route:

$$BW = \sum_{k=1}^{N_r} B_k, \tag{3}$$

where N_r is the number of possible intermediate routes and is the B_k is the bandwidth use for the kth route.

3.5. Handoff Delay due to the Rerouting Procedures

To calculate the handoff delay, T_h, we sum the time to send the signaling messages for each scheme. The overall handoff rerouting delay is:

$$T_h = m * \sum_{l=1}^{N_s} T_l, \tag{4}$$

where m is the number of times a new route is determined, N_s is the number of steps between the first handoff initiation and the completion of the final route as illustrated in Figures 1 and 2, and T_l is the time to execute step l of the handoff rerouting procedure.

The delay to deliver a signaling message, M_l, is:

$$M_l = (\alpha_l + \beta_l + \gamma_l) * H_l, \tag{5}$$

where α_l is the transmission time, β_l is the propagation time, and γ_l is the processing time for the control message in signaling step l. H_l is the number of hops between the switches involved in the signaling exchange. The *transmission time*, α_l, is computed:

$$\alpha_l = \frac{S_l}{B_l}, \tag{6}$$

where S_l is the size of the control message in bits, and B_l is the bit rate of the link on which the message is sent.

Using the message delivery time, M_l, in (5), we can calculate T_l from 4 separately for the wireline and wireless links. For the wireline links, the signaling execution time is the same as the message delivery time:

$$T_l = M_l. \tag{7}$$

For the wireless link, we must consider the probability of wireless link failure. Let n_f be the number of wireless link failures. Then for the wireless link, T_l is:

$$T_l = \sum_{n_f=0}^{\infty} T_l(n_f) * Prob\{n_f \text{failures and 1 success}\} \tag{8}$$

Let the waiting time to determine that the message was lost be equal to the time to send one message. Then $T_l(n_f)$ is:

$$T_l(n_f) = M_l + n_f * (2 * M_l), \tag{9}$$

and T_l becomes:

$$\begin{aligned} T_l &= \sum_{n_f=0}^{\infty} [M_l + n_f * (2 * M_l)] * \\ &\quad Prob\{n_f \text{failures and 1 success}\} \qquad (10) \\ &= M_l + (2 * M_l) * \\ &\quad \sum_{n_f=0}^{\infty} n_f * Prob\{n_f \text{failures and 1 success}\} \qquad (11) \end{aligned}$$

Given a probability q that the wireless link fails, we use the techniques in [6] to calculate the sum in (11). Then the signaling execution time over the wireless link, T_l is:

$$T_l = M_l + (2 * M_l) * \frac{q}{1-q}. \tag{12}$$

Thus, the time to execute the signaling, T_l, for each step l of the handoff rerouting protocol is:

$$T_l = \begin{cases} M_l, & \text{for the wireline links} \\ M_l * \frac{1+q}{1-q}, & \text{for the wireless links} \end{cases} \tag{13}$$

and is used to determine the overall handoff delay for the handoff rerouting scheme as in 4.

3.6. Buffer Requirements

To find the buffer requirement, Q, per connection, we determine the steps for which the handoff process must buffer data traffic. Then we calculate the amount of buffer space used during the execution of those steps:

$$Q = B * \sum_{l=L}^{N_s} T_l, \tag{14}$$

where B is the bit rate of the channel, L is the first step in the handoff algorithm where buffering is required, N_s is the number of steps until the buffer is released, and T_l is the time to execute signaling step l, as found in (13).

Each of the above calculations is carried out for the new scheme as well as the related handoff rerouting schemes. In the next section, we explore the numerical results for the protocol comparison.

4. Numerical Results

The system parameters used to compare the protocols are found in Table 1. The processing delays are considered different for routing and computation. Routing processing delay refers to computing the next route, while computation processing delay refers to special computations, such as finding the anchor switches or calculating a new group in a multicast protocol [4]. Note that since our new scheme relies on the source switch to manage the handoff rerouting, and since the rerouting is performed differently for loss sensitive and delay sensistive traffic, the results of our new scheme are labeled *Source-Loss* and *Source-Delay*.

Parameter	Value
Message Size (S, (6))	400 bits
Signaling Channel Bandwidth, (B_l, (6))	64 Kbps
Wireless Data Channel Bandwidth (B_k, (3))	64 Kbps
Wireline Data Channel Bandwidth (B_k, (3))	155 Kbps
Wireless Propagation Delay, (β, (5))	5 μsec
Wireline Propagation Delay (β, (5))	50 μsec
Route Processing Delay (γ, (5))	0.5 $m$sec
Computation Processing Delay (γ, (5))	5 $m$sec
Static Route Length (2)	2 hops

Table 1. System Parameters

4.1. Final Route Length

We the calculated final route length, R_f, (1), for the following two cases:

1. The MTs move only one hop, and then rest,
2. The MTs move three consecutive hops and then rest.

Figure 4 shows the results. The new technique for loss sensitive traffic, Source-Loss, and the chaining technique [7, 5] both resulted in an optimal route, equal to the end-to-end distance. (The two results lie on top of each other.) The worst performance for final route length was shown by the multicast technique [4], which was especially affected by its forwarding process. For the three hop case, the new Source-Delay scheme showed an improvement in final route length performance compared to the static/dynamic anchored rerouting method [9, 2, 8].

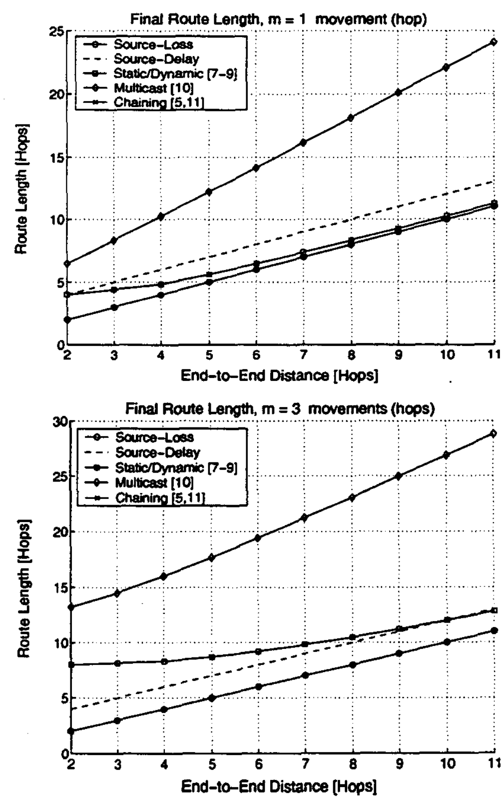

Figure 4. Final Route Length from the New Source to the New Destination

4.2. Summed Route Length

Figure 5 shows the results of the summed route lengths, R_s, (2), for $m = 1$ and $m = 3$ hops by each MT. We found that for one hop movement, the multicast [4] and chaining [7, 5] techniques have a larger summed route length than the new Source-Loss and Source-Delay results, illustrating the affect of repeated connection set up/tear down. For three hops, the summed route length of all of the schemes more than doubles—except for the new technique for the Source-Loss results, which consistently obtain one route calculation at the optimal route length.

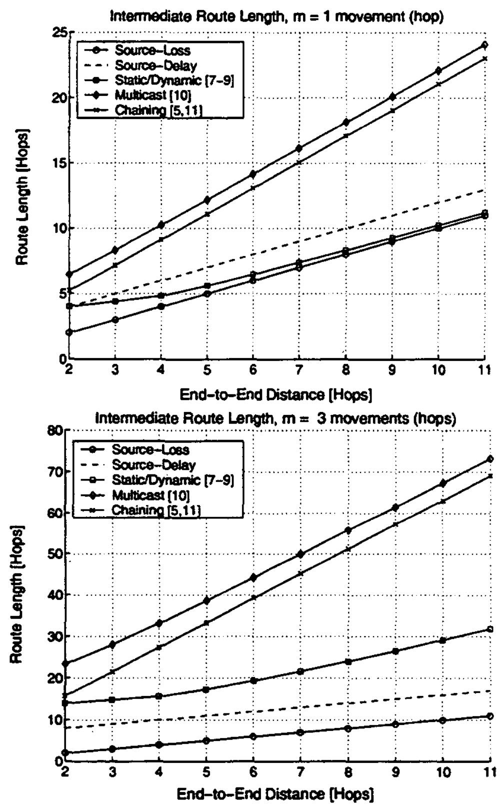

Figure 5. Summed Length of the Intermediate Routes

4.3. Bandwidth Consumption

The bandwidth consumption due to the intermediate routes, BW, (3), is illustrated in Figure 6. For the multicast [4] and chaining [7, 5] examples, the bandwidth use increases $1\frac{1}{2}$ to 2 times each consecutive MT movement. The Source-Delay performance has the same bandwidth consumption as the chaining technique [7, 5], since both schemes use the initial path extension. The static/dynamic technique [9, 2, 8] shows better performance. However, since the new Source-Loss technique uses only one route calculation, no additional bandwidth is set up and torn down for loss sensitive connections.

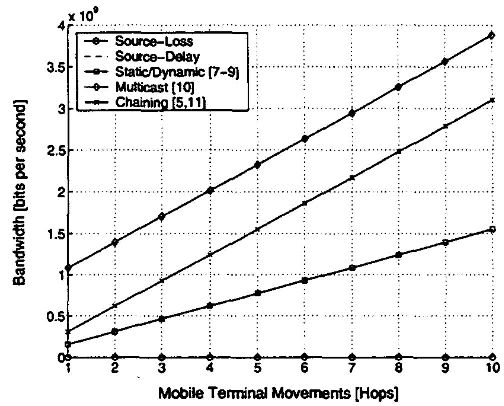

Figure 6. Bandwidth Consumption due to Intermediate Routes

4.4. Handoff Delay

Next, the handoff delay due to multiple consecutive rerouting operations, T_h, (4), is shown in Figure 7. The Source-Loss technique and the chaining technique [7, 5] are twice as slow as the static/dynamic scheme [9, 2, 8] and the multicast scheme [4], due to the repeated source-to-destination signaling procedures. However, the route extension in the Source-Delay technique reduces the handoff delay for delay sensitive traffic in the new scheme.

4.5. Buffering Required per Connection

Finally, the buffer requirement for one MT movement, Q, (14), is shown in Figure 8. As expected, negligible buffering is used by the new technique for delay sensitive traffic, Source-Delay. (The path alignment buffering is not included in the calculation for any of the schemes.) The static/dynamic scheme [9, 2, 8] shows the next best performance, followed by chaining scheme [7, 5] and the new scheme for loss sensitive traffic, Source-Loss. However, when multiple connections are present, the buffer space will be multiplied for each connection for the other schemes, but will be multiplied only for the loss sensitive traffic for the new scheme. For example, for two connections per MT,

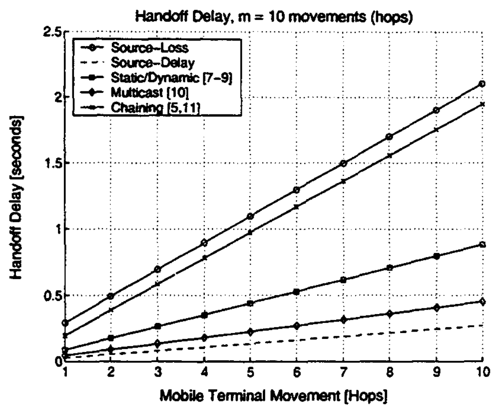

Figure 7. Handoff Delay due to Rerouting

one delay sensitive connection and one loss sensitive connection, the related schemes double the buffer requirement in Figure 8 while the new technique will remain the same as in Figure 8.

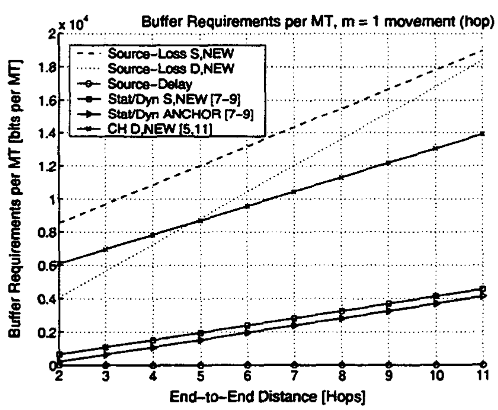

Figure 8. Buffering Requirements per Connection

5. Conclusion

Future wireless networks must be able to support multimedia traffic over mobile connections. In this paper, we presented a handoff rerouting scheme for connection-oriented networks that supports multimedia traffic over connections with multiple mobile endpoints. We used a common framework to compare the new scheme with related work and found that the new technique exhibited a superior performance with respect to route length and bandwidth consumption. Our scheme achieves an optimal route for loss sensitive traffic at the expense of handoff delay. However, our scheme allows delay sensitive connections to experience the smallest handoff delay by taking advantage of the extension method. The buffer requirement of the new scheme is competitive when multiple connections requiring different QoS constraints are considered.

References

[1] I. F. Akyildiz, J. McNair, J. Ho, H. Uzunalioğlu, and W. Wang. Mobility Management in Next Generation Wireless Systems. *Proceedings of the IEEE*, 87(8), August 1999.

[2] S. Biswas and A. Hopper. A Representative Based Architecture for Handling Mobility in Connection Oriented Radio Networks. In *Proc. of the International Conference on Universal Personal Communications (ICUPC)*, November 1995.

[3] B. Edmaier, J. Eberspächer, W. Fischer, and A. Klug. Alignment Server for Hitless Path-switching in ATM Networks. In *Proc. of XV Int. Switching Symp. (ISS '95)*, pages 403–407, Berlin, Germany, April 1995.

[4] R. Ghai and S. Singh. An Architecture and Communication Protocol for Picocellular Networks. *IEEE Personal Communications Magazine*, 1:36–47, Third Quarter 1994.

[5] IS-41 Revision C. Cellular Radio-Telecommunications Intersystem Operations. Technical report, EIA/TIA, 1995.

[6] L. Kleinrock. *Queueing Systems*. Wiley, 1975.

[7] B. Rajagopalan. Mobility Management in Integrated Wireless-ATM Networks. *ACM-Baltzer Journal of Mobile Networks and Applications (MONET)*, 1(3):273–285, 1996.

[8] M. Veeraraghavan, M. Karol, and K. Eng. Mobility and Connection Management in a Wireless ATM LAN. *IEEE Journal on Selected Areas in Communications*, 15(1):50–68, January 1997.

[9] R. Yuan, S. Biswas, L. J. French, J. Li, and D. Raychaudhuri. A Signaling and Control Architecture for Mobility Support. *ACM-Baltzer Journal of Mobile Networks and Applications (MONET)*, 1(3):287–298, January 1996.