Chapter 1

PERFORMANCE EVALUATION OF THE HIERARCHICAL HANDOFF SCHEME FOR WIRELESS ATM NETWORKS

*Nadjia Kara, &Jean Conan, #André Girard

**ECE Departement, Ecole polytechnique, Montreal, Canada, kara@inrs-telecom.quebec.ca*

&ECE Departement, Ecole polytechnique, Montreal, QC, Canada, conan@com.polymtl.ca

#INRS-Telecom, Université du Québec, Québec, QC, Canada. andre@inrs-telecom.quebec.ca

Key words: Handoff, Wireless ATM.

Abstract: One of the most important issues in wireless networks is handoff. Recently, a hierarchical protocol has been proposed to manage handoff in wireless ATM networks based on the partial rerouting strategy and the virtual connection tree concept. In this paper, a discrete time model is proposed to estimate the connection bandwidth and the connection handoff delay of this scheme. We propose to compare the hierarchical protocol with the Akyol and Cox handoff protocol. Numerical results illustrate that for the future wireless ATM networks the hierarchical model gives lower connection bandwidth and lower connection handoff delay than those of Akyol and Cox model.

1. INTRODUCTION

Two features are dominating in the development of future communication networks: ATM technology with its high capacity and service capabilities and guaranteed Quality of Service (QoS) on the one hand, and support of mobility on the other hand. The combination of these features is wireless ATM. Wireless ATM is expected to play a significant role in the next generation of wireless broadband networks (Raychaudhuri & Wilson, 1994; The ATM Forum, 1998; Bush et al., 1997) and since no specific standards are yet specified for wireless ATM, this topic is an active research area. Many aspects of wireless ATM have been studied and

mobility including connection rerouting to support handoff is one of the most important ones and is the focus of our work.

Figure 1. Partial rerouting scheme

Several rerouting protocols have been proposed to simplify handoff management (Acompora & Naghshineh, 1994; Akyol & Cox, 1997; Veerarghavan et al., 1996; Worng & Leung, 1998). One of these is the partial rerouting scheme based on a partial connection re-establishment (Akyol & Cox, 1997). In this case, a switch along the original connection is selected to setup a portion of connection to the target switch to which the new base station is connected (Fig. 1). The switch selected to perform this function is called the crossover switch. This protocol gives an efficient use of bandwidth by setting up a new optimal path. Despite this efficient management of bandwidth, the time to compute this optimal path may exceed the constraint of handoff delay.

Another protocol is the virtual connection tree scheme, where resources are preallocated in the area surrounding the base station to which a mobile terminal is connected (Acompora & Naghshineh, 1994). The tree root is a specialized ATM switch with mobility management capability (Fig. 2). This protocol ensures a fast handoff but the bandwidth use is inefficient.

Figure 2. Virtual connection tree scheme

To solve the delay and bandwidth problems, we have proposed in (Kara et al., 1999) a new handoff protocol called hierarchical handoff. In this model, we use trees regrouping base stations and we regroup all of these trees into an arborescence structure connected to the root node. This structure implies two kinds of handoffs : the intra-subnetwork handoff and the inter-subnetworks handoff. The intra-subnetworks handoff is managed by the tree root nodes of the subnetwork. Since in this case the resources are always available, the handoff may occur without delay (for example in figure 3 the tree with root node G for handoff from P_1 to P_2). The inter-subnetworks handoff is managed by using the partial rerouting procedure but only between the root nodes of the subnetworks, for example between C and J for the handoff from P_2 to P_3 (Fig. 3). This handoff scheme allows a distributed mobility management since each subnetwork root node manages the mobility for a set of base stations independently of other root nodes of the subnetwork. The signaling traffic in the network is thus reduced and this reduction is very important for multimedia networks. We expect that this handoff scheme will give a handoff delay and bandwidth smaller than those of the Akyol and Cox handoff scheme.

In this paper, we estimate the bandwidth and handoff delay of the hierarchical model and we compare this performance criteria with those of the Akyol and Cox model. This paper is organized as follows. In section 2, the hierarchical handoff scheme is described and performance analysis of this model is presented where the performance criteria are the bandwidth and handoff delay. In section 3, we describe the performance analysis of the Akyol and Cox handoff scheme by estimating the bandwidth and the handoff delay. In section 4, we present numerical results and the comparison between the two handoff schemes. Conclusion is given in section 5.

Figure 3. Hierarchical handoff scheme

2. ANALYTICAL MODEL OF THE HIERARCHICAL HANDOFF SCHEME

The hierarchical handoff protocol is a combination of the partial rerouting protocol and the virtual connection tree protocol. When the mobile moves within the subnetwork, intra-subnetwork handoff may occur. In this case, the connection is maintained by the virtual connection tree, and the connection rerouting duration is very short. If the mobile moves outside the subnetwork, then a combined rerouting which consists of a partial rerouting followed by a virtual connection tree rerouting, is performed to maintain the connection. A discrete time model is developed in (Chan & Leung, 1998), for the two-phase handoff scheme, for evaluating the bandwidth of three different path optimization schemes. We propose to adapt and to generalize this model to evaluate the bandwidth and the handoff delay for the hierarchical model. Events that may occur at each time $i=1,2,\dots$ are: 1) partial rerouting; 2) call termination. The intra-subnetwork handoff is not included because it does not generate signaling traffic. At each time i , we define:

- λ the probability that there is an inter-subnetwork handoff (and thus a partial rerouting);
- μ the probability that the call terminates.

Thus, inter-subnetwork handoff is modeled as a Bernoulli process. The call duration has a geometric distribution with mean $1/\mu$. For each mobile connection, we define :

- L_h the number of links between the root node of the subnetwork to which the mobile belongs and the destination during the call duration.

L_h is a random variable with a general distribution and with mean $\overline{L_h}$. We assume that the number of links of the new connection (rerouted connection) is a random variable with the same mean $\overline{L_h}$.

Figure 4. Time diagram

The performance measures are the expected bandwidth per call and the expected handoff delay per call. The expected bandwidth per call is the network bandwidth needed to support a mobile connection, and is the sum of the link bandwidth and the signaling bandwidth due to inter-subnetworks handoff. Likewise, the expected handoff delay per call is the network handoff delay to support a mobile connection, and is the sum of the link delay (delay of user's bandwidth reservation) and the signaling delay due to the inter-subnetworks handoff. Let:

- B be the bandwidth of one call per unit time on a given link;
- D be the delay of the bandwidth reservation of one call per unit time on a given link;
- B_{RP} be the signaling bandwidth for each path rerouting;
- D_{RP} be the signaling delay for each path rerouting.

The parameters B_{RP} and D_{RP} include the bandwidth and the delay respectively for : 1) locating the handoff switch; 2) setting up the new portion of connection; 3) terminating the old portion of connection; and 4) updating the connection server about the status of the existing route.

In the following subsections the problem of determining the expected bandwidth per call and the expected handoff delay per call for the hierarchical handoff scheme is addressed.

By using a discrete time model, we will estimate the expected handoff delay per call denoted $E[D_h]$, and the expected bandwidth per call $E[B_h]$.

The time diagram of figure 4 subdivides the call duration into time intervals. In each time interval $[i, i+1]$ a handoff may occur. In each interval, let:

- D_i be the delay of the bandwidth reservation for one call during time interval $[i, i+1]$ called delay;
- D_i^l be the delay of bandwidth reservation on the links of one call during time interval $[i, i+1]$ called link delay;
- D_i^s be the signaling delay due to handoff of one call during time interval $[i, i+1]$;
- B_i be the bandwidth for one call during time interval $[i, i+1]$ called bandwidth;
- B_i^l be the bandwidth reserved on the links of one call during time interval $[i, i+1]$ called link bandwidth;
- B_i^s be the signaling bandwidth due to handoff of one call during time interval $[i, i+1]$.

Let n be the call termination time. D_i , D_i^l , and D_i^s , are related by the following equation:

$$D_i = D_i^l + D_i^s \quad (1)$$

Likewise B_i , B_i^l , and B_i^s are related by :

$$B_i = B_i^l + B_i^s \quad (2)$$

We define a random variable indicator

$$I_i = \begin{cases} 1 & \text{if an inter-subnetwork handoff occurs at time } i, \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Since at each time interval the partial rerouting is modeled as a Bernoulli random variable with probability 1, we have $E[I_i]=1$.

2.1 Expected handoff delay per call

The signaling delay of the partial rerouting due to mobility at time i is

$$D_i^s = D_{RP} I_i \quad (4)$$

the expected of D^s is

$$E[D_i^s] = \lambda D_{RP} \quad (5)$$

and the link delay between time interval $]i, i + 1]$ is

$$D_i^l = L_h D I_i \quad (6)$$

The expectation of D_i^l is

$$E[D_i^l] = \lambda \overline{L_h} D \quad (7)$$

and the expected delay per call is

$$\begin{aligned} E[D_h] &= E[E[D_h / N]] = \\ &= \sum_{n=1}^{\infty} E\left[\sum_{i=0}^{n-1} D_i | N = n\right] P(N = n) = \sum_{n=1}^{\infty} \sum_{i=0}^{n-1} E[(D_i^l + D_i^s)] P(N = n) = \\ &= \lambda \frac{1-\mu}{\mu} \overline{L_h} D + \lambda \frac{1-\mu}{\mu} D_{RP} \end{aligned} \quad (8)$$

2.2 Expected bandwidth per call

The signaling bandwidth for partial rerouting due to mobility at time i is

$$B_i^s = B_{RP} I_i \quad (9)$$

the expectation of B^s is

$$E[B_i^s] = \lambda B_{RP} \quad (10)$$

and the link bandwidth between time interval $]i, i + 1]$ is

$$B_i^l = B_{i-1}^l (1 - I_i) + L_h B I_i \quad (11)$$

Taking the expectation of the link bandwidth :

$$E[B_i^l] = (1 - \lambda) E[B_{i-1}^l] + \lambda \overline{L_h} B \quad (12)$$

the expectation bandwidth per call is :

$$\begin{aligned}
E[B_h] &= E \left[E \left[\sum_{i=0}^{n-1} B_i / N \right] \right] = \\
\sum_{n=1}^{\infty} E \left[\sum_{i=0}^{n-1} B_i | N = n \right] P(N = n) &= \sum_{n=1}^{\infty} \sum_{i=0}^{n-1} E[B_i^l + B_i^s] P(N = n) = \quad (13) \\
\frac{1-\mu}{\mu} \overline{L_h} B + \lambda \frac{1-\mu}{\mu} B_{RP}
\end{aligned}$$

To this bandwidth we add a fixed bandwidth B_{hs} due to the resource reservation in the subnetworks.

$$B_{hs} = L_{hs}(B + B_{RP}) \quad (14)$$

Where L_{hs} is the number of links in the subnetworks. Since the expected bandwidth per call is

$$E[B_h] = \frac{1-\mu}{\mu} \overline{L_h} B + L_{hs} B + \lambda \frac{1-\mu}{\mu} B_{RP} + L_{hs} B_{RP} \quad (15)$$

the expected handoff delay per call and the expected bandwidth per call of the hierarchical handoff scheme have been estimated. We will now estimate the expected handoff delay per call and the expected bandwidth per call of Akyol and Cox handoff scheme.

3. ADAPTED ANALYTICAL MODEL TO THE AKYOL AND COX HANDOFF SCHEME

The Akyol and Cox handoff protocol is based on the partial rerouting (Akyol & Cox, 1997). At each inter-switch handoff, a partial rerouting is performed. The performance analysis in (Akyol & Cox, 1997) estimates the number of signaling messages and the signaling bandwidth of the protocol. In this section, we adapt the analytical model of the hierarchical handoff scheme to Akyol and Cox handoff scheme to derive the expected handoff delay per call and the expected bandwidth per call and then to compare the hierarchical handoff scheme with the Akyol and Cox handoff scheme.

The rerouting is modeled as the hierarchical rerouting model with the following definitions :

- λ is the probability that there is an inter-switch handoff (and thus a partial rerouting);

– L_S is the number of links between the source and destination.

L_S is a random variable with a general distribution and with mean L_S . We suppose that the number of connection links after the partial rerouting is a random variable with the same mean L_S .

We define a random variable indicator :

$$J_i = \begin{cases} 1 & \text{if an inter-switch handoff occurs at time } i \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

At each time interval the partial rerouting is modeled as a Bernoulli random variable with probability λ and $E[J_i] = \lambda$.

In this case, we can compute the expected handoff delay per call and the expected bandwidth per call.

3.1 Expected handoff delay per call

In this case, following the same steps as established for the hierarchical handoff scheme, we arrive at the following results.

The signaling delay of the partial rerouting due to inter-switch handoff at time i is

$$D_i^s = D_{RP} J_i \quad (17)$$

the expected signaling delay due to mobility at time i is

$$E[D_i^s] = \lambda D_{RP} \quad (18)$$

The link delay between time interval $]i, i+1]$ is

$$D_i' = L_S D J_i \quad (19)$$

the expected link delay between time interval $]i, i+1]$ is

$$E[D_i'] = \lambda \overline{L_S} D \quad (20)$$

The expected handoff delay per call is

$$E[D_S] = \frac{\lambda(1-\mu)}{\mu} (\overline{L_S} D + D_{RP}) \quad (21)$$

3.2 Expected bandwidth per call

The signaling bandwidth for partial rerouting due to inter-switch handoff at time i is

$$B_i^s = B_{RP} J_i \quad (22)$$

the expected signaling bandwidth due to mobility at time i is

$$E[B_i^s] = \lambda B_{RP} \quad (23)$$

The link bandwidth between time interval $]i, i+1]$ is

$$B_i^l = L_S B J_i \quad (24)$$

the expected link bandwidth between time interval $]i, i+1]$ is

$$E[B_i^l] = (1 - \lambda) E[B_{i-1}^l] + \lambda \overline{L_S} B \quad (25)$$

Then, the expected bandwidth per call is

$$E[B_S] = \frac{1 - \mu}{\mu} (\overline{L_S} B + \lambda B_{RP}) \quad (26)$$

In the following section, we present the numerical results of each model and compare the results of the hierarchical handoff technique with those of the Akyol and Cox handoff scheme.

4. NUMERICAL RESULTS

The expected handoff delay in equations (8) and (21) are normalized by the signaling delay for partial rerouting D_{RP} . The expected bandwidth in equations (15) and (26) are normalized by the signaling bandwidth for partial rerouting B_{RP} .

If the time unit is second then for $\mu=0.03$ the mean call duration is equal to 33.33 sec. The mean call duration is 1–2 minutes for the voice traffic and 15–20 minutes for data traffic, this mean durations correspond to $\mu 0.03$ and in this case, we have the following results.

Figures 5 and 6 illustrate the expected handoff delay per call as a function of λ of the hierarchical handoff and Akyol and Cox schemes. The expected handoff delay of each handoff scheme increase as λ increases. The performance comparisons between these two handoff protocols show that when D/D_{RP} increases, the hierarchical model gives an expected handoff delay lower than the one of Akyol and Cox model. This means that the higher the mobility in network, the lower the handoff delay of the hierarchical handoff scheme. When λ increases, the expected bandwidth of the hierarchical and the Akyol and Cox models increases (Fig. 7 and 8). Moreover, when B/B_{RP} is large, the expected bandwidth of the hierarchical model is lower than that of the Akyol and Cox model. This means that for networks with connections with large bandwidths the expected bandwidth per call for the hierarchical model is lower than the one of the Akyol and Cox model. We can derive a rule in which cases one would use the hierarchical or the Akyol and Cox protocols :

- 1 If the network operator must carry only one kind of traffic (e.g. voice) then :
 - a If the QoS (handoff delay) is more important than the bandwidth then we suggest to use the hierarchical handoff scheme.
 - b If the bandwidth is more important than the QoS then we suggest to use the Akyol and Cox handoff Scheme.
- 2 If the network operator must carry multimedia traffic then we suggest to use the hierarchical handoff scheme.

5. CONCLUSION

In this paper, an analytical model to evaluate the expected handoff delay per call and the expected bandwidth per call of the hierarchical handoff scheme has been proposed. This model has been adapted to the Akyol and Cox handoff scheme in order to be able to compare the expected handoff delay and the expected bandwidth per call of the Akyol and Cox and the hierarchical handoff schemes. Numerical results show that the hierarchical model yields a lower expected handoff delay and a lower expected bandwidth per call than those of Akyol and Cox for the multimedia wireless ATM networks when the mobility (λ) increases.

Figure 5. Expected handoff delay per call for the hierarchical and the Akyol and Cox models
 $(D/D_{RP} = 0.5, \bar{L}_S = 5, \bar{L}_h = 1, L_{hs} = 60)$

Figure 6. Expected handoff delay per call for the hierarchical and the Akyol and Cox models
 $(D/D_{RP} = 1, \bar{L}_S = 5, \bar{L}_h = 1, L_{hs} = 60)$

Figure 7. Expected bandwidth per call for the hierarchical and the Akyol and Cox models
 $(B/B_{RP} = 0.5, \bar{L}_S = 5, \bar{L}_h = 1, L_{hs} = 60)$

Figure 8. Expected bandwidth per call for the hierarchical and Akyol and Cox models

$$(B/B_{RP} = 1, \overline{L}_S = 5, \overline{L}_h = 1, L_{hs} = 60)$$

REFERENCES

- [1] Acampora and Naghshineh, 1994 Acampora, A. and Naghshineh, M. (1994). An architecture and methodology for mobile-executed handoff in cellular ATM networks. *IEEE Journal on Selected Areas in Communications*, 12(8):1365–1375.
- [2] Akyol and Cox, 1997 Akyol, B. and Cox, D. (1997). Signaling alternatives in a wireless ATM network. *IEEE Journal on Selected Areas in Communications*, 15(1):35–49.
- [3] Bush et al., 1997 Bush, S., Jagannath, S., Evans, J., Frost, V., Sanchez, R., Minden, G., and Shanmugan, K. (1997). A control and management network for wireless ATM systems. *Wireless Networks*, 3:267–283.
- [4] Chan and Leung, 1998 Chan, W. W. H. and Leung, V. (1998). Performance evaluations of path optimization schemes for inter-switch handoffs in Wireless ATM networks. *Proceedings of ACM/IEEE MobiCom '98, Dallas, Texas*, pages 242–251.
- [5] Kara et al., 1999 Kara, N., Conan, J., and Girard, A. (1999). Handoff management in wireless ATM networks. *The Eleventh International Conference on Wireless Communications, Calgary*, pages 48–54.
- [6] Raychaudhuri and Wilson, 1994 Raychaudhuri, D. and Wilson, N. (1994). ATM-based transport architecture for multiservices wireless personal communication networks. *IEEE Journal on Selected Areas in Communications*, 12(8):140–1414.
- [7] The ATM Forum Wireless ATM Working Group, 1998 The ATM Forum Wireless ATM Working Group (1998). Baseline Text for Wireless ATM Specifications. *BTDM-WATM-01.07*.
- [8] Veeraraghavan et al., 1996 Veeraraghavan, M., Karol, M., and Eng, K. (1996). A combined handoff scheme for mobile ATM networks. *ATM Forum Contribution 96-1700*.
- [9] Wong and Leung, 1998 Wong, W. and Leung, V. (1998). A path optimization signaling protocol for inter-switch handoff in wireless ATM networks. *Proceedings of 1st International Workshop on Wireless Mobile ATM Implementation (WMATM'98), Hangzhou, China*, pages 92–98.