A Two-State Markovian Queueing Model to Describe Handoff Behavior of a VBR Connection in a Wireless ATM Network

Chih-Yang Tsai and Jin-Fu Chang

Abstract—**Handoff behavior induced by user mobility is a unique phenomenon existing in wireless mobile communication networks. A connection once established may have to undergo a number of handoffs before it is hung up or dropped. A two-state Markovian queueing model is constructed in this paper to understand the performance of a VBR call in a wireless ATM network when handoffs are taken into consideration. The performance measures examined are the mean and jitter of an ATM cell. The model is extended to tri-state if call termination due to handoff failure is further incorporated in the model.**

I. Introduction

MOST of us have witnessed in just the past few years how the booming wireless communication service have fastly penetrated to impact and enrich our works and lives. Internet in the air shall soon become a reality. It is not the purpose of this section of our paper to review the progress and the advancement of this rapidly growing wireless communication technology and project where it should be pushed forward. Rather we shall concentrate on one technical issue that is not only crucial but also unique in this new communication paradigm. That is the handoff phenomenon due to user mobility in a wireless mobile communication network.

Handoff has traditionally been divided into two basic types. Hard handoff are frequently employed in FDMA (frequency division multiple access) and TDMA (time division multiple access) while soft handoffs are oriented for the CDMA (code division multiple access) systems.

Handoff issue has so far attracted considerable attention.Past works on handoff include [1]-[15] to treat problems such as user mobility models, traffic model, signalling scheme, adptation to the TCP/IP or ATM networks, bandwidth provisioning , and etc.

The asynchronous transfer mode (ATM) has been in recent years a mainstream technique to support broadband integrated servcies. In this integrated environment services are divided into the following four types according to their QoS (quality of service) requirement : constant bit rate (CBR), variable bit rate (VBR), available bit rate (ABR) and unspecified bit rate (UBR). With the emergence of wireless technology, the extension of the ATM to the mobile domain is now dubbed the wireless ATM (WATM). This is why there have been also papers (e.g. [4], [7], [16]-[17]) to look into the handoff problems in the ATM environment. In this paper we elect to examine the behavior of a VBR connection that may encounter in its entire course of communication a random number of handoffs. Our main interest lies in the delay performance, including mean and variance, of the VBR connection.

Chih-Yang Tsai is with Quanta Computer Inc., 188, Wen Hwa 2nd Rd.,Kuei shan, Tao Yuan, Taiwan 333, email: cyrus.tsai@quantatw.com

Jin-Fu Chang is with Department of Electrical Engineering, National Chi Nan University, Puli, Nantou, Taiwan 545, email: jfchang@ncnu.edu.tw (Corresponding author)

This work was done while the authors were with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan 10617.

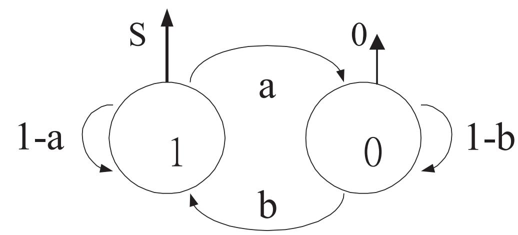


Fig. 1. A two state user mobility model.

When travelling in a cell, a VBR source is permitted to transmit at a rate coordinated through a medium access control (MAC) protocol. While moving from one cell to a neighboring cell, the VBR source may be temporarily blocked from transmitting in the course of a hard handoff or maybe still transmitting at a lower rate during a soft handoff. It is the purpose of this paper to understand the delay performance of a VBR connection when taken the handoff behavior of either kind into account.

To reflect user mobility, we introduce the following 2×2 mobility matrix M, and represent it graphically in Fig. 1. In this figure, 0 and 1 are used to represent "handoff" and "residence" state, respectively.

$$M = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix},$$

In this paper we assume a frame-based MAC protocol for the coordination of channel access among different users. TDMA is such a typical frame-based protocol. For frame-based protocols, we further assume that state change occurs only at frame boundaries. In other words, cell residence and handoff transition time are both geometrically distributed. State change at frame boundaries is well justified since resource allocation is also handled on the frame-by-frame basis.

The rest of this paper is organized as follows. Sec. II treats the hard handoff problem while soft handoff problem is dealt in Sec. III. A two-state Markovian model is used to describe the behavior of a VBR connection from call set-up till termination. In Sec. IV we also consider the situation when call termination due to handoff failure occurs. Finally , concluding remarks are given in Sec. V.

II. Hard Handoff

To understand the effect of hard handoff on the uplink delay performance of a VBR connection, we examine the queueing behavior of the connection at frame boundaries. To this end, Fig.1 can be employed to model a server with interrupted services caused by handoff. When the server comes to state 1 (i.e. , the mobile is in residence state), the server serves the queues at a constant rate, i.e., a fixed number of packets/cells are transmitted during each frame as long as the mobile remains in the residence state. The problem regarding how many packets can be sent per frame is decided when the connection is established.

In this paper, we focus on this scenario of fixed rate of service. When the mobile begins to undertake a handoff procedure (i.e., when the mobile switches to state 0 from state 1) , the service is temporarily halted.

To tackle this problem, we proceed in three phases. First, we acquire the queue length of the VBR connection at the beginning of a frame. Second, using the queue lengths obtained in phase 1 we obtain the queue length at the arrival instant of a representative cell whose delay we wish to find. Third, based on the queue length at the arrival instant of a cell, we analyse its delay behavior.

A. Phase 1: Queue length at frame boundaries

For the VBR connection we are examining, we use X_n to denote the queue length , i.e., number of packets or cells, of the VBR connection at the beginning of the n-th frame. Likewise, we use A_n and S_n to denote the number of cells that have arrived and departed respectively, in the n-th frame.Then

$$X_{n+1} = (X_n + A_n - S_n)^+, \quad (1)$$

where (x)$^+$= x if x ≥ 0, and 0 if x<0.

To deterrmine the steady-state statictics of X, we need to first find the transition probability from (X_n, I_n, J_n) to $(X_{n+1}, I_{n+1}, J_{n+1})$ where I_n and J_n are used to denote the server state and the phase of the two-state MMPP arrival process, respectively.

When $I_n = 1$ then $S_n = s$ (the number of time slots appropriated for the VBR connection when it was admitted) and $S_n = 0$ when $I_n = 0$. The transition probability $\Pr ob[X_{n+1} = x_{n+1}, I_{n+1} = i_{n+1}, J_{n+1} = j_{n+1} | X_n = x_n, I_n = i_n, J_n = j_n]$ can be obtained in a straight-forward manner and is summerized in [18].

After the individual transition probabilities are obtained, we place them into an M/G/1 type transition matrix $T = [t_{ij}]$of the following form

$$\begin{bmatrix} B_0 & B_1 & B_2 & B_3 & B_4 & B_5 & . & . & . \\ A_0 & A_1 & A_2 & A_3 & A_4 & A_5 & . & . & . \\ 0 & A_0 & A_1 & A_2 & A_3 & A_4 & . & . & . \\ 0 & 0 & A_0 & A_1 & A_2 & A_3 & . & . & . \\ 0 & 0 & 0 & A_0 & A_1 & A_2 & . & . & . \\ 0 & 0 & 0 & 0 & A_0 & A_1 & . & . & . \\ 0 & 0 & 0 & 0 & 0 & A_0 & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & . & . & . \\ . & . & . & . & . & . & . & . & . \end{bmatrix}, \quad (2)$$

where

$$B_0 = \begin{bmatrix} b_{0,0} & \cdots & b_{0,m-1} \\ \vdots & & \vdots \\ b_{m-1,0} & \cdots & b_{m-1,m-1} \end{bmatrix},$$

$$B_j = \begin{bmatrix} b_{0,mj} & \cdots & b_{0,mj+m-1} \\ \vdots & & \vdots \\ b_{m-1,mj} & \cdots & b_{m-1,mj+m-1} \end{bmatrix},$$

$$A_0 = \begin{bmatrix} a_0 & a_1 & \cdots & a_{m-1} \\ 0 & a_0 & \cdots & a_{m-2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_0 \end{bmatrix},$$

$$A_j = \begin{bmatrix} a_{mj} & \cdots & a_{mj+m-1} \\ \vdots & & \vdots \\ a_{mj+m-1} & \cdots & a_{mj} \end{bmatrix}, for \quad j \geq 1,$$

so that the algorithmic approach developed in [19] [20] can be employed to find $\Pr ob(X = x)$ at steady state and its moments.

A.1 Phase 2: Queue length at the time of arrival of a cell

Let the time of arrival of an arbitrary cell be t such that $t \in [0, T_f)$, where T_f is the length of a frame.

We use $X_t(z)$ to denote the PGF (probability generating function) of queue length distribution at time t. Then

$$X_t(z) = X(z) \cdot A^*(z,t) \quad (3)$$

where $A^*(z,t)$ tells the number of cells that have arrived ahead of the cell we are examining.

Due to a two-state MMPP [21]

$$A^*(z,t) = \exp(Q - (1-z)\Lambda)t, \quad (4)$$

where

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix}, \quad and \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

are the parameter matrices of the two-state MMPP.

Since a cell may arrive at any time within a frame randomly, we may obtain the z-transform of $\widehat{X}$, the queue length at any arbitary cell arrival point, as follows,

$$\widehat{X}(z) = \frac{1}{T_F}\int_0^{T_F} X_t(z) \; dt = \frac{1}{T_F}\int_0^{T_F} X(z) \cdot A^*(z,t) \; dt. \quad (5)$$

The first and second moment of $\widehat{X}$, i.e $\widehat{X}^1(1)$ and $\widehat{X}^2(1)$, can be found through the technique developed in the MMPP cookbook[21], and details can be found.

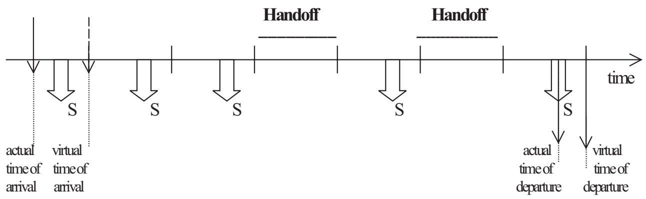


Fig. 2. A snapshot to illustrate the decomposition of cell delay.

B. Phase: 3 Delay decomposition

We use Fig. 2 to illustrate how the delay of a cell is accounted. In the treatment of delay analysis, we assume like many other studies that a cell can commence its transmission no sooner than the beginning of the next frame after the arrival of this cell has occured.

In Fig. 2 we show the actual arrival and departure time of a cell whose delay we wish to investigate. Since a cell may arrive at any time within a frame and its departure may also occur anywhere within a later frame, we may very well restrict the arrival and departure of this cell to occur at the end of the frame of actual arrival and departure for the purpose of delay calculation. For this, we also show the virtual arrival and departure time of the cell in Fig. 2.

Having paved the way, the delay of the cell we are examining becomes tangible. Upon the arrival of this cell, it sees $\widehat{X}$ cells of (5) already ahead. Let us call a frame within which the VBR connection may send up to s cells an ON frame.On the contrary, if the VBR conection is blocked from transmitting cells due to handoff, then that frame is an OFF frame.

The $\widehat{X}+1$ cells shall consume a total of $R = \left\lceil (\widehat{X}+1)\ /\ s \right\rceil^{+}$ ON frames where $\lceil x \rceil^{+}$ is the smallest integer larger than or equal to x.

As shown in Fig. 2 two ON frames maybe separated by zero, one, two, or $\cdots$ OFF/handoff frames. That is, the gap between two ON frames could be of length zero, one, two, ... frames. We use the random variable G to denote the size of a gap, i.e., $G = 0, 1, 2, \cdots$ From the model in Fig. 1 we have

$$\Pr ob[G = g] = 1 - a, \;\; as \;\; g = 0, \tag{6}$$

and

$$\Pr ob[G = g] = ab(1-b)^{g-1}, \;\; as \;\; g \geq 1, \tag{7}$$

with PGF

$$G(z) = \frac{(1-a)(1-z) + bz}{(1-z) + bz}.$$

We use D to denote the delay of the cell we are examining, then

$$D = (G+1) \times R \;\; (frame). \tag{8}$$

The following result has been obtained in [22]

$$R(z^s) = z\left(\frac{1-z^s}{s \times z^{s-1}} \times \sum_{i=0}^{s-1} \frac{w^i \widehat{X}(w^i z)}{1 - w^i z}\right) \tag{9}$$

where

$$w = \exp(j2\pi/s)$$

From the above equation we further obtain

$$\begin{aligned} E\{R\} &= \frac{dR(z)}{dz}\Big|_{z=1} \\ &= 1 + \frac{1}{s}\left[\widehat{X}(1) - \frac{s-1}{2} - \sum_{i=0}^{s-1} \frac{w^i \widehat{X}(w^i)}{1-w^i}\right] \;, \end{aligned} \tag{10}$$

$$\begin{aligned} E\{R(R-1)\} &= \frac{d^2R(z)}{dz^2}\Big|_{z=1} \\ &= 2E\{R\} + \frac{1}{s}\Big\{\frac{1}{s}\Big[\widehat{X}(1) - \frac{s-1}{2} - \\ &\quad \sum_{i=0}^{s-1} \frac{w^i \widehat{X}(w^i)}{1-w^i}\Big](1-s) - \\ &\quad \frac{1}{s}\Big[\frac{1}{3}(1-s^2) + (s-1)\widehat{X}^1(1) - \widehat{X}^2(1) + \\ &\quad \sum_{i=1}^{s-1} \frac{(1-w^i)w^{2i}\widehat{X}^1(w^i) + w^{2i}\widehat{X}(w^i)}{(1-w^i)^2}\Big]\Big\}. \end{aligned} \tag{11}$$

Using (6)-(11), we can obtain the mean and variance of delay described in (8).

Let us show some numerical examples using the result we have obtained so far. We consider a TDMA system in which each frame contains 50 time slots that collectively last 20 ms.

Let the VBR soure be modeled by a two-state MMPP having the following parameter matrices

$$Q = \begin{bmatrix} -0.0075 & 0.0075 \\ 0.0025 & -0.0025 \end{bmatrix}, \Lambda = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.1 \end{bmatrix}.$$

This corresponds to an average cell generation rate of $\bar{\lambda} = \pi\lambda = [0.25, 0.75][0.3, 0.1]^T = 0.15$ cells/slot. This VBR source has a burstiness of $0.3/0.15 = 2$, that resembles a video signal[23].

Table 1 gives three (a,b) pairs of Fig. 1 to be used in our numericals to reflect three different kinds of mobility.

TABLE I

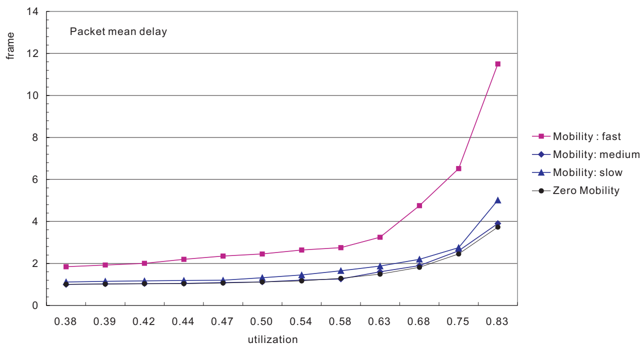
THREE TYPES OF MOBILITY.

mobility	a	b
slow	0.0002	0.2
medium	0.002	0.2
fast	0.2	0.2

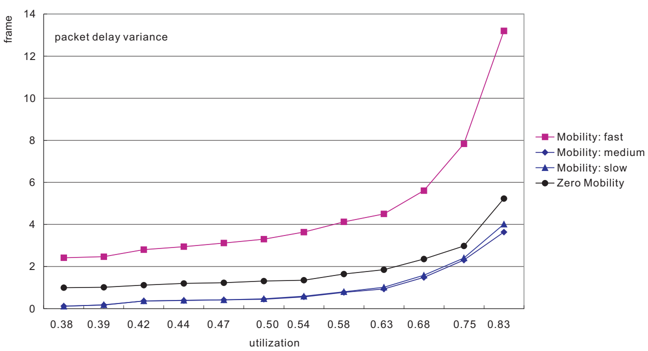
Under the same value of b =0.2, take a = 0.002 as an example. This corresponds to that a handoff occurs on the average of every 500 frames or 10 sec.

If cells are of size 10 to 100 meters, then mobiles move at a speed of 1m/sec to 10 m/sec and is considered “slow”. That b = 0.2 says every handoff takes an average of 5 frames or 100ms to complete.

Fig. 3 plots delay mean and variance versus utilization/load for the three types of mobility specified in Table1. Here utilization is defined to be $\bar{\lambda} \cdot T_F/s$. We wish to point out that the accuracy of these results have been verified by computer simulations. In Fig. 3 we also plot an additional case with zero mobility. The trend that both mean and varianve increase as utilization grows is reasonable. We further observe that among the four cases, fast mobility exhibits the largest mean and variance, than medium , slow and zero mobility. As a matter of fact, the results of slow and zero mobility almost coincide.



(a)



(b)

Fig. 3. Packet mean delay and delay variance with different mobility parameter sets.

III. Soft Handoff

Refer to Fig. 1 again, a mobile now may still transmit at a nonzero rate during handoff. Let s_0 and s_1 be the rate that the mobile transmits at state 0 and 1, respectively, of Fig 1.

It is reasonable to assume that $s_1 \geq s_0 \geq 0$.Clearly, the problem we treated in Sec. 2 is a special case of the problem here with $s_0 = 0$ and $s_1 = s$.

Phases 1 and 2 of our previous analysis in Sec. 2 can be extended in a straightforward manner. For example, S_n in (1) now becomes s_0 or s_1. But phase 3 becomes a lot more complicated . We shall in the following present only phase 3.

In Sec. 2, frames are divided into two kinds: ON and OFF. But here we divide them into simply 0 and 1 with rate s_0 and s_1. Suppose $\widehat{X}+1$ cells are to be sent in n_0 0-frames and n_1 1-frames, then $D = n_0 + n_1$frmaes and n_0, n_1 must satisfy

$$n_0 \cdot s_0 + n_1 \cdot s_1 \geq \widehat{X} + 10.2cm;$$

but

$$(n_0 - 1) \cdot s_0 + n_1 \cdot s_1 < \widehat{X} + 10.2cm, \tag{12}$$

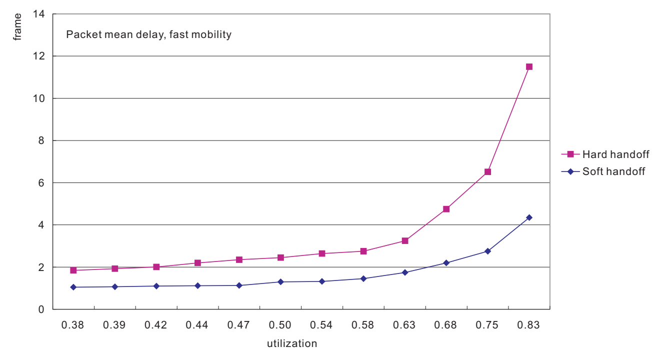
if the last frame is a 0-frame; and

$$n_0 \cdot s_0 + (n_1 - 1) \cdot s_1 > \widehat{X} + 10.2cm, \tag{13}$$

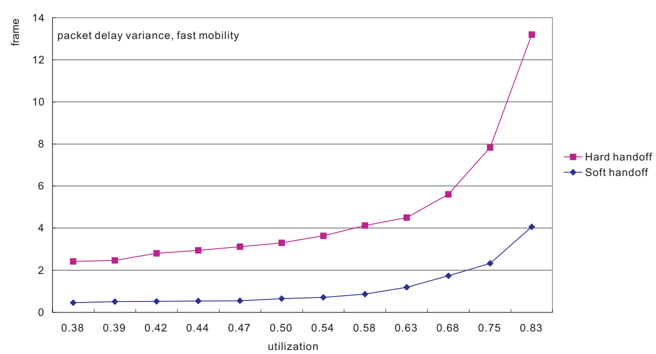
if the last frame is a 1-frame.

The (n_0, n_1) problem we posted in the above is a difficult combinational problem. We proceed here with an alternative approximation. We employ the model we used in Sec. 2 with $s = (b \cdot s_0 + a \cdot s_1)/(a+b)$ and use the results we obtained there to find the cell mean delay and delay variance. The accuracy is well supported by the computer simulations we have conducted in [18].

Regarding numerical examples we use again the system we employed in Sec 2 and compare the difference between hard and soft handoff. Figs. 4-6 show delay mean and variance versus utilization for the slow, medium, and fast mobility specified in Table 1. In these figures, we observe the reasonable result that soft handoff outperforms hard handoff and the difference widens as mobility increases.



(a)



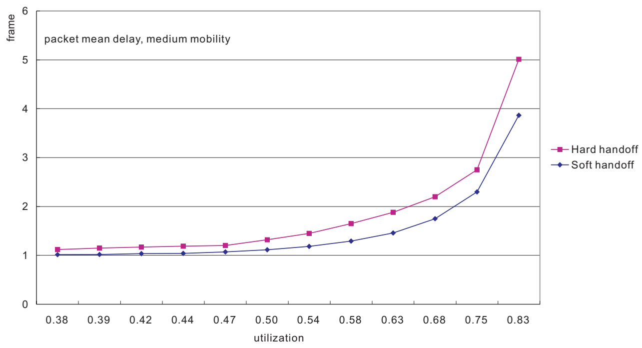
(b)

Fig. 4. Packet mean delay and delay variance comparisons of hard handoff and soft handoff with fast mobility.

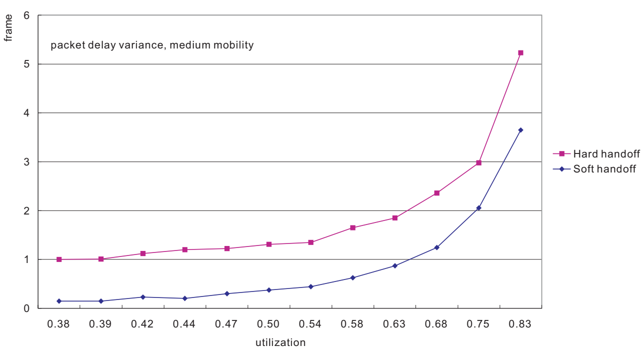
IV. Hanodff Failure

It is a well recognized fact that a mobile wireless call maybe terminated due to handoff failure. To reflect this fact, the 2-state model is expanded to a 3-state model in Fig. 7.

In Fig. 7(a) the handoff failure state represented by -1 is an absorption state, so that the transition from handoff state 0 to -1 represents a call termination. But what happens in practice is when a call has not been completed, the subscriber normally redials to get a new connection. This is represented by the state transition diagram in Fig 7(b). In this paper we investigate the

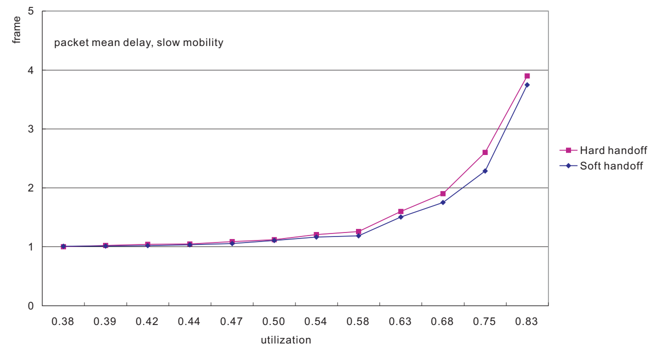


(a)

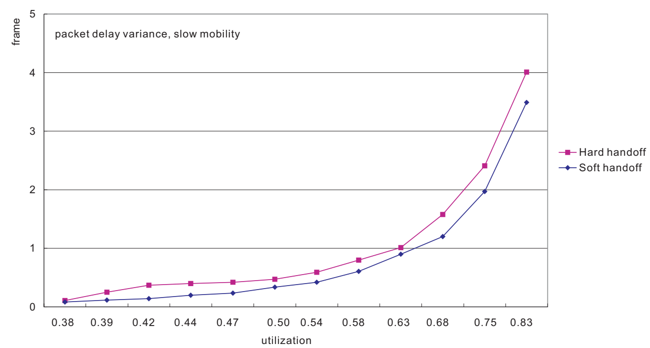


(b)

Fig. 5. Packet mean delay and delay variance comparisons of hard handoff and soft handoff with medium mobility.



(a)



(b)

Fig. 6. Packet mean delay and delay variance comparisons of hard handoff and soft handoff with slow mobility.

situation that corresponds to Fig. 7(b).

Although the mobility model is now changed to three states. The analysis resembles Sec.2 and 3. Thus we shall present only a numerical example that uses the same set of parameters. The result is now plotted in Fig. 8. In this figure, we consider two handoff failure probabilities 10^{-3} and 10^{-6}. We also include the result of no handoff failure in Fig. 8. First, we observe that no handoff failure exhibits the best performance among these three curves. Second, a handoff failure probability of 10^{-6} behaves nearly as no handoff failure. This is also reasonable. Third , only when the handoff failure probability gets to the order of 10^{-3} we start to see a noticeable difference.

V. Conclusion

We have in this paper employed a 2-state or 3-state mobility model to understand the effect of handoff on the delay performance of a VBR call in a WATM environment. We use a 2-state MMPP to describe the traffic generated by the VBR source. Three kinds of mobility - slow, medium, fast are considered in numerical experiments. In both hard handoff and soft handoff slow mobility exhibits the best performance, i.e., smallest delay mean and varianve, then medium and fast. Soft handoff is seen to perform better than hard handoff. When handoff failure is taken into consideration, then performanve is deteriorated from hard handoff. But performance degradation can be negleted if handoff failure probability does not exceed 10^{-3}.

The analysis in this section can be extended to other types of traffic, e. g. , CBR. The only difference is the model to describe the traffic source. Although our analysis is conducted for the uplink, delay behavior on the downlink may be similarly treated.

References

[1] M. M. Zonoozi, P. Dassanayake, and M. Faulkner, *Mobility Modeling and Channel Holding Time Distribution in Cellular Mobile Communication Systems*, IEEE ICC'95.

[2] K. Sohraby, *Mobility Calculus in Cellular Wireless Networks*, IEEE ICC'96.

[3] F. Khan, and D. Zeghlache, *Effect of Cell Residence Time Distribution on The Performance of Cellular Mobile Networks*, IEEE VTC'97.

[4] L. H. Tang, and H. D. Hughes, *A Handoff Scheme for Wireless ATM Networks*, IEEE ICUPC'96.

[5] M. Litzenburger, H. Bakker, and W. Schodl, *Handover in a Broadband Cellular ATM Access System*, IEEE VTC'98.

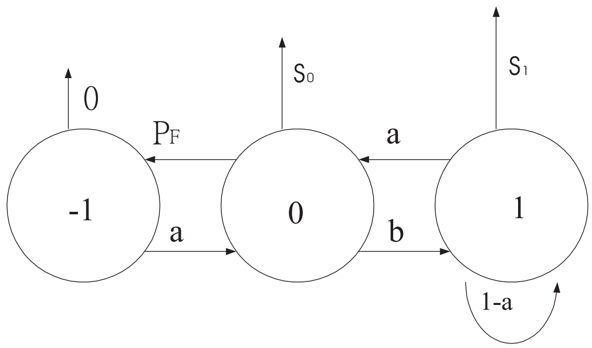
[6] R. Ramjee, T. F. LaPorta, J. Kurose, and D. Towsley, *Performance Evaluation of Connection Rerouting Schemes for ATM-Based Wireless Networks*, IEEE/ACM Trans.Networking, vol. 6, no. 3, pp. 249-261, June 1998.

[7] S. Shiokawa, and S. Tasaka, *Bandwidth Allocation for Wireless ATM Considering Priorities of Component Calls*, IEEE Globecom'99.

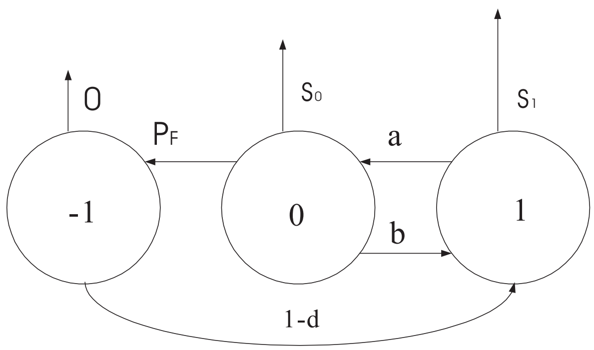
[8] M. M. Zonoozi, and P. Dassanayake, *Handover Delay and Hysteresis Margin in Microcells and Macrocells*, IEEE PIMRC'97.

[9] C. Vargas, M. V. Hegde, and M. Naraghi-Pour, *Blocking Effects of Mobility and Reservations in Wireless Networks*, IEEE ICC'98.

[10] S. Tekinay, and B. Jabbari, *A Measurement-Based Prioritization Scheme*

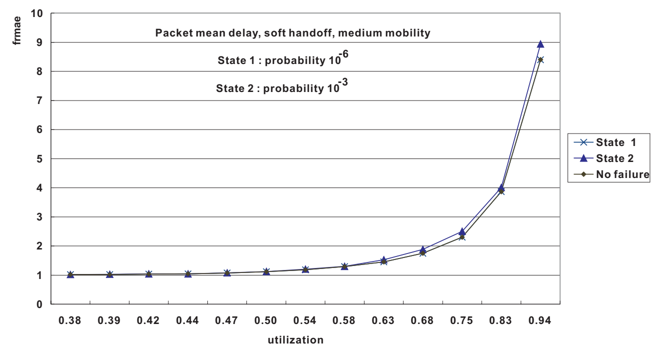


(a)

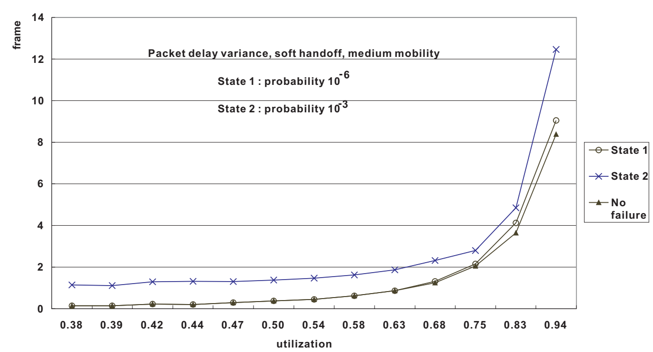


(b)

Fig. 7. (a) and (b). A 3-state mobility model that includes handoff failure.



(a)



(b)

Fig. 8. (Packet mean delay and delay variance with different call failure probability in soft handoff with medium mobility.

for Handovers in Mobile Cellular Networks, IEEE J. Select. Areas Commun., vol. 10, no. 8, pp. 1343-1350, October 1992.

[11] D. Calin, and D. Zeghlache, *Performance and Handoff Analysis of an Integrated Voice-Data Cellular System*, IEEE PIMRC'97.

[12] D. McMillan, *Delay Analysis of a Cellular Mobile Priority Queueing System*, IEEE/ACM Trans.Networking., vol. 3, no. 3, pp. 310-319, June 1995.

[13] W. Li, and A. S. Alfa, *A PCS Network with Correlated Arrival Process and Splitted-Rating Channels*, IEEE J.Select.Areas Commun., vol. 17, no. 7, pp. 1318-1325, July 1999.

[14] C. Oliveira, J. B. Kim, and T. Suda, *An Adaptive Bandwidth Reservation Scheme for High-Speed Multimedia Wireless Networks*, IEEE J. Select. Areas Commun., vol. 16, no. 6, August 1998.

[15] Q. -A. Zeng, K. Mukumoto, and A. Fukuda, *Performance Analysis of Two-level Priority Reservation Handoff Scheme in Mobile Cellular Radio Systems*, IEEE VTC'97.

[16] B. A. Akyol, and D. C. Cox, *Rerouting for Handoff in a Wireless ATM Network*, IEEE ICUPC'96.

[17] B. G. Marchent, and M. J. McTiffin, *Handover and Macro Diversity for 3rd Generation Mobile Systems within ATM Fixed Networks*, IEEE Globecom'96.

[18] C. Y. Tsai, *The Impact of User Mobility on Delay Variation of Frame-Based Multiple Access Protocol*, Master Thesis, Graduate Institute of Communication Engineering, National Taiwan University, June 2000.

[19] M. F. Neuts, *Structured Stochastic Matrix of M/G/1 Type and their Applications*, Marcel Dekkar Inc., 1989.

[20] D. M. Lucantoni and V. Ramaswami, *Efficient Algorithms for Solving the Non-linear Matrix Equations Arising in Phase type Queues*, Stochastic Models, vol. 1(1), pp. 29-51, 1985.

[21] W. Fischer, and K. Meier-Hellstern, *The Markov-Modulated Poisson Process (MMPP) Cookbook*, Performance Evaluation 18, pp. 149-171, 1993.

[22] H. Bruneel, *Message Delay in TDMA Channels with Contiguous Output*, IEEE Trans. Common., vol. COM-34, pp. 681-684, July 1986.

[23] R. Handel, M. N. Huber, and S. Schroder, *ATM Networks, Concepts, Protocols, Applications*, Addison-Wesley, 1994.