Impact of PCS Handoff Response Time

Yi-Bing Lin, *Senior Member, IEEE*

***Abstract*— In a personal communications services (PCS) network, the network delay for a handoff request is limited by a timeout period. If the network fails to respond within the timeout period, the handoff call is forced terminated. We study the effect of the network response time on the performance (the call incompletion probability) of a PCS network. Our study indicates that at a small offered load, the network response time has a significant effect on the call incompletion probability. We also observe that the effect of the network response time is more significant if the mobile residence time distribution at a cell has a smaller variance.**

***Index Terms*— Forced termination, handoff, network response, personal communications.**

I. Introduction

IN A *personal communications services* (PCS) network, the service area is populated with a large number of base stations (BS's) with each providing coverage in its vicinity. When a call arrives at the *coverage area* of a base station or *cell*, the destination (or the originating) mobile phone is connected if a channel is available. Otherwise, the call is blocked (this is referred to as a *new call blocking*). When a communicating mobile phone moves from one cell to another, the channel in the old BS is released, and a channel is required in the new BS. This process is called *handoff*. In a perfect environment (where handoff occurs instantly), a handoff call continues if an idle channel exists at the new BS. In a practical environment, there is an handoff time delay before the mobile phone connects to the new BS. The handoff time delay is constrained by a timeout period. If the delay is longer than the timeout period, the handoff call is forced terminated. The impact of network response time on handoff performance has not been studied in the literature. We investigate this issue. Specifically, we study the *call incompletion probability* or the probability that a call is not completed due to a new call blocking or a forced termination. Assume that the network response times have a density function f_N, and the timeout period is τ. Define α as the probability that the network fails to complete the handoff procedure within the timeout period. Then

$$\alpha = \Pr[t > \tau] = \int_{t=\tau}^{\infty} f_N(t)dt.$$

Manuscript received May 19, 1997. The associate editor coordinating the review of this letter and approving it for publication was Prof. A. K. Elhakeem. This work was supported in part by National Science Council, R.O.C., under Contract NSC-87-2213-E-009-013.

The author is with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. (e-mail: liny@csie.nctu.edu.tw).

Publisher Item Identifier S 1089-7798(97)08954-0.

In this letter, the network response time is characterized by the probability α that the network fails to complete the handoff procedure within the timeout period (for a particular network response time distribution, α can be computed by $\Pr[t > \tau]$). In some specific PCS networks, α is limited to be less than 0.1.

II. An Analytic Model

We propose an analytic model to study the effect of the probability α that the network fails to complete the handoff procedure within the timeout period on the call incompletion probability. The following assumptions are used in the model.

- The call arrivals to/from a mobile phone are a Poisson process. The net new call arrival rate to a cell is λ_o.
- The mobile residence times t_m in a cell have a general distribution with the density function $f_m(t)$ and the Laplace transform $f^*_{\mathbf{m}}(s)$. The expected mobile residence time is $E[t_m] = 1/\eta$.
- The call holding time t_c is exponentially distributed with the mean $1/\mu$.

The output measures are:

- λ_h handoff call arrival rate to a cell;
- p_b probability that no idle channel is available when a new/handoff call arrives;
- p_o $= p_b$—new call blocking probability;
- p_f $= 1-(1-\alpha)(1-p_b) = 1-(1-\alpha)(1-p_o)$—forced termination probability or the probability that a handoff call is blocked. Note that a handoff call is forced terminated if the network handoff response time is too long (with probability α) or no channel is available (with probability p_b);
- p_{nc} call incompletion probability.

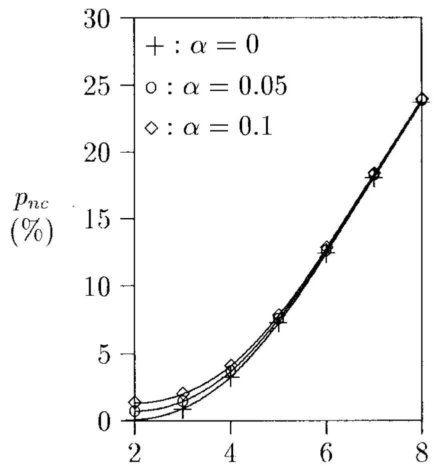
We modify the derivation of λ_h in [1] by considering the probability α that the network fails to process handoff within the timeout period. Consider a communicating mobile phone in a cell. Let π_1 (π_2) be the probability that a new call (handoff call) of the mobile phone is not completed before the mobile phone moves out of the cell, then from [1], we have

$$\pi_1 = \frac{\eta}{\mu}\left[1 - f^*_{\mathbf{m}}(\mu)\right] \quad \text{and} \quad \pi_2 = f^*_{\mathbf{m}}(\mu).$$

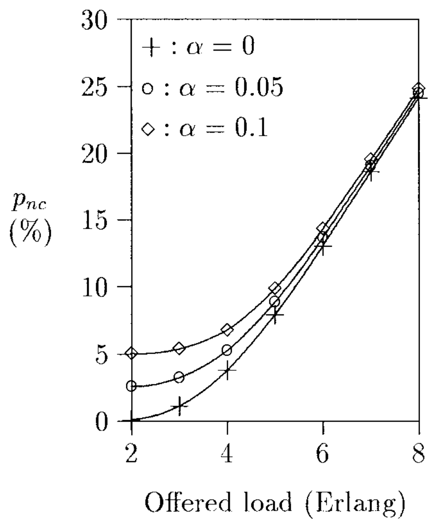
Consider a homogeneous PCS network structure where the rate of the handoff calls flow in a cell is the rate of the handoff calls flow out of the cell. Then we have

$$\lambda_h = \lambda_o(1-p_b)\pi_1 + \lambda_h(1-p_f)\pi_2 \tag{1}$$

$$= \frac{\eta(1-p_o)[1-f^*_{\mathbf{m}}(\mu)]\lambda_o}{\mu[1-(1-\alpha)(1-p_o)f^*_{\mathbf{m}}(\mu)]} \tag{2}$$



(a)



(b)

Fig. 1. The interaction between α and η ($c = 8$). (a) $\lambda = 0.008$, $\eta = 0.125\mu$. (b) $\lambda = 0.008$, $\eta = 0.5\mu$.

Equation (1) implies that a handoff call overflows from a cell to its neighbors in two cases.

- A new call is not blocked (with probability $1 - p_b = 1 - p_o$) at the cell and the call is not completed before the mobile phone leaves the cell (with probability π_1).
- A handoff call is not forced terminated (with probability $1 - p_f = (1-\alpha)(1-p_b) = (1-\alpha)(1-p_o)$) at the cell and the call is not completed before the mobile phone leaves the cell (with probability π_2).

The probability p_{nc} can be derived from (2):

$$p_{nc} = p_o + \frac{\lambda_h p_f}{\lambda_o} \tag{3}$$

$$= p_o + \frac{\eta(1-p_o)[1-f^*_{\mathbf{m}}(\mu)][1-(1-\alpha)(1-p_o)]}{\mu[1-(1-\alpha)(1-p_o)f^*_{\mathbf{m}}(\mu)]}. \tag{4}$$

Equation (3) is interpreted as follows. In a period Δt, there are $\lambda_o \Delta t$ new call arrivals to a cell. These new calls generate $\lambda_h \Delta t$ handoff calls. Among these new/handoff calls, the number of blocked calls is $p_o \lambda_o \Delta t + p_f \lambda_h \Delta t$. Thus p_{nc} is expressed as (3). Note that $p_{nc} \neq p_o + p_f$ because an incomplete call may successfully make several handoffs before it is forced terminated.

The expected *dwell time* (i.e., the channel occupancy time) $E[t_{do}]$ for a new call is derived in [2]

$$E[t_{do}] = \frac{1}{\mu} - \frac{\eta}{\mu^2}\left[1 - f^*_{\mathbf{m}}(\mu)\right]. \tag{5}$$

The expected dwell time for a handoff call is

$$E[t_{dh}] = \frac{1}{\mu}\left[1 - f^*_{\mathbf{m}}(\mu)\right]. \tag{6}$$

The probability p_b is computed by using the following iterative algorithm and (2), (5), and (6).

The Iterative Algorithm

Input Parameters: λ_o (the new call arrival rate), μ (the call completion rate), c (the number of channels in a BS), and $f_m(t_m)$ (the mobile residence time density function).

Output Measures: λ_h (the handoff call arrival rate), $p_b = p_o$ (the probability that no idle channel is available when a call arrives or the new call blocking probability), and p_{nc} (the call incompletion probability).

Step 1: Select an initial value for λ_h.

Step 2: Compute p_o (to be elaborated).

Step 3: $\lambda_{h,old} \leftarrow \lambda_h$.

Step 4: Compute λ_h by using (2):

$$\lambda_h = \frac{\eta(1-p_o)[1-f^*_{\mathbf{m}}(\mu)]\lambda_o}{\mu[1-(1-\alpha)(1-p_o)f^*_{\mathbf{m}}(\mu)]}.$$

Step 5: If $|\lambda_h - \lambda_{h,old}| > \delta\lambda_h$ then go to Step 2. Otherwise, go to Step 6. Note that δ is a pre-defined value.

Step 6: The values for λ_h and p_o converge. Compute p_{nc} by using (4):

$$p_{nc} = p_o + \frac{\eta(1-p_o)[1-f^*_{\mathbf{m}}(\mu)][1-(1-\alpha)(1-p_o)]}{\mu[1-(1-\alpha)(1-p_o)f^*_{\mathbf{m}}(\mu)]}.$$

In Step 2, the system under study can be modeled as an M/G/c/c queue. By using (5) and (6), the net traffic to a cell is

$$\rho = \lambda_o E[t_{do}] + \lambda_h(1-\alpha)E[t_{dh}]$$
$$= \frac{\lambda_o}{\mu}\left\{1 - \frac{\eta(p_o + \alpha - \alpha p_o)[1-f^*_{\mathbf{m}}(\mu)]}{\mu[1-(1-\alpha)(1-p_o)f^*_{\mathbf{m}}(\mu)]}\right\}.$$

Since the blocking probability for an M/G/c/c queue is the same as an M/M/c/c queue [3],

$$p_o = p_b = B(\rho, c) = \frac{(\rho^c/c!)}{\sum_{i=0}^{c}(\rho^i/i!)}$$

where $B(\rho, c)$ is the Erlang loss equation.

III. Discussion and Conclusions

We discuss the impact of α on p_{nc} based on the analytic model in the previous section. Assume that the mobile residence times in a cell have a Gamma distribution [4]. Depending upon the values of the parameters, the Gamma distributions can be shaped to represent many distributions as well as shaped to fit sets of measured data determined by the mobile phone speed, moving direction, the size of a cell, and the signal propagation.

For any $\gamma > 0$, a Gamma density function is

$$f_\gamma(t) = \frac{\beta^\gamma t^{\gamma-1} e^{-\beta t}}{\Gamma(\gamma)} \quad \text{where} \quad \Gamma(\gamma) = \int_{\tau=0}^{\infty} e^{-\tau}\tau^{\gamma-1} d\tau \tag{7}$$

where β is called the *scale* parameter, and γ is called the *shape* parameter. The mean value of the Gamma residence time distribution is $\frac{1}{\eta} = \frac{\gamma}{\beta}$ and its Laplace transform is

$$f^*_{\mathbf{m}}(s) = \left(\frac{\gamma\eta}{s+\gamma\eta}\right)^\gamma.$$

Our analysis indicates the following.

1) Interaction Between α and the Mobile Phone Mobility: Fig. 1 shows that the effect of α on p_{nc} is more significant for a larger mobility η. Intuitively, if $\eta = 0$ (the mobile phones are stationary), then every conversation is completed before the mobile phone leaves the cell, and α does not have any effect on p_{nc}.

2) Interaction Between α and the Offered Load: Fig. 1 indicates that the effect of α is more significant for a small offered load than a large offered load. In other words, it is important to reduce the network response time for a small offered load. At a large offered load, reducing the network response time does not improve the performance. We also observe (not shown in Fig. 1) that the effect of the network response time is more significant if the mobile residence time distribution at a cell has a smaller variance.

Acknowledgment

The author would like to thank K.C. Chen for his contribution to this work.

References

[1] Y.-B. Lin, S. Mohan, and A. Noerpel, "Queueing priority channel assignment strategies for handoff and initial access for a PCS network," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 704–712, Aug. 1994.

[2] Y. Fang, I. Chlamtac, and Y.-B. Lin, "Call performance for a PCS network," *IEEE J.Select. Areas Commun.*, to be published.

[3] L. Kleinrock, *Queueing Systems: Volume I – Theory*. New York: Wiley, 1976.

[4] N. L. Johnson, *Continuous Univariate Distributions-1*. New York: Wiley, 1970.