Effects of Mobility on Mean Packet Delay over Integrated Services Wireless Networks

J. Padilla
UPB, Colombia
jpadilla@upbbga.edu.co

J. Paradells
UPC, Spain
teljpa@entel.upc.es

A. Rodríguez
UPB, Colombia
alvaro.rodriguez@ieee.org

Abstract

Next generation cellular networks will be based on packet switching and will have QoS support based on architectures such as Integrated Services. Then, to calculate QoS parameters such as packet delay will be very important. In order reach to this goal, in this paper we propose a model for an Integrated Services wireless access router. Our model describes the user mobility effect on the mean packet delay over such routers under two different CAC schemes. Besides, we apply this model to evaluate the mobility effect over two QoS support approaches, the Standard IntServ and a new proposal for QoS support based on Integrated Services named IntServ6. [1]

1. Mobility effect over the mean packet delay

Our goal is to model the network layer response independently from the physical layer. Thus, our model does not take into account physical layer issues. As the main QoS parameter for Integrated Services networks is the mean packet delay, we modeled this issue and its relationship with mobility parameters such as the new calls number and the handoff calls number. At the same time, these parameters depend on the admission control schemes.

Until now, several approaches have worked on mobility effects over Call Admission Control Schemes (CAC schemes) in cellular networks. Some works, have studied some CAC particular schemes and assumes that the channel holding times of new calls and handoff calls are independent and identically distributed (i.i.d.) according to an exponential distribution . But field data show that channel holding time does not necessarily obey the exponential distribution assumption. In a series of studies [1-3], Fang and others show that the new call channel holding time and handoff call channel holding time may have different distributions and different average values. However, a new approximation approach is proposed in [3] to avoid solving the large set of flow equations. In this approach, new calls and handoff calls are assumed independent and exponentially distributed but with different average values. Analysis and simulation results show that the approximation approach in [3] is more realistic and it performs much better than the traditional approach. Then, our study is based on this Fang's proposal.

As we have mentioned above, these works are focused on mobility effect over CAC schemes and they obtained expressions over the calls performance level. Thus, they obtained parameters as new call blocking probability and handoff dropping probability. However, next generation cellular networks will be based on packet switching and will have QoS support based on architectures as Integrated Services. Then, to calculate QoS parameters as packet delay will be very important. Recently, in [4], a proposal to model packet delay and overflow probability was performed for a very simple wireless base station modeled as a general multiplexer. However, it is necessary to obtain more complex models that allow network designers to describe these issues in a better way.

In order to achieve this goal, our approach introduces a more complex model and describes the user mobility effect over the mean packet delay on an Integrated Services packet switching cellular access router under two different CAC schemes: the new call bounding scheme and the cutoff priority scheme. Also, in this paper we apply this model to evaluate the mobility effect over two QoS approaches, the standard IntServ approach [5] and the IntServ6 proposal [6], which is a new proposal for QoS support based on Integrated Services.

We introduce in Fig. 1 our wireless access router model. We can observe that an Integrated Services wireless Access Router is mainly composed of two parts; the first part is the reservations manager, which is composed by the admission control scheme and the reservations setup agent.

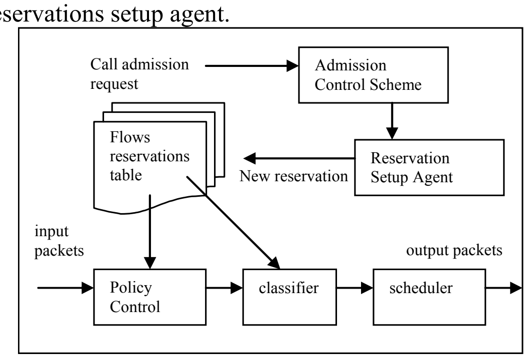

Figure 1. Wireless access router model.

[1] This work was supported by ERDF and the Spanish government through project TIC2003-01748.

The second part performs the packet forwarding process. It is composed by the policies control module, the classification process and the scheduling process. As we can see in fig. 1, these two main parts are related across the Flows Reservation table. Then, when a new call is accepted, the number of flows on the Flows Reservation table will increase in one flow. Thus, when a packet arrives, it should be classified. In other words, it should be found on the Flows reservations table. Finally, it should be scheduled over the correspondent output link. In our model, we suppose that the policy control module delay is zero. This aspect will be studied in future works. Based on the described model, we have developed a mathematical model which details several issues. The first issue is the relationship between the number of flows (or calls) that the Flows reservations table holds and the new calls traffic intensity and handoff traffic intensity. In order to accomplish this, we modeled two admission control schemes: the Cuttoff Priority scheme and the New Call bounding scheme. The second modeled issue is the mean packet burst delay on the classifier. This parameter depends on the number of flows (or calls) that the Reservations table contains. A third modeled issue is the mean packet delay for the scheduling process. Later, we will explain all models.

1.1. Mobility model and admission control schemes

To model our approach, we have used the Fang Mobility Model described on [3]. This model assumes that the arrival processes of new call and handoff calls are Poisson type. The channel holding time for new and handoff calls is assumed independent and exponentially distributed but with different average values. Let λ_n, λ_h, $1/\mu_n$ y $1/\mu_h$ denote the new calls arrival rate and the handoff calls arrival rate, and the average channel holding times of new calls and handoff calls, respectively. Let $\rho_n=\lambda_n/\mu_n$ and $\rho_h=\lambda_h/\mu_h$. Let C denote the total number of channels in a cell. Using the Fang mobility model, call level performance, e.g., new call blocking probability (P_{nb}) and handoff dropping probability (P_{hb}), of QoS provisioning schemes for several CAC can be derived. The CAC schemes modeled in [3] and used in our approach are explained below.

1.1.1. New Call Bounding Scheme. In this scheme, we limit the admission of new calls into the wireless networks. The scheme works as follows: if the number of new calls in a cell exceeds a threshold when a new call arrives, the new call will be blocked; otherwise it will be admitted. A handoff call is rejected only when all channels in the cell are used up.

Let n_1 denotes the number of new calls initiated in the cell and n_2 denotes the number of handoff calls in the cell. Also, let K be the threshold for the new calls. Expressions for new call blocking probability ($P_{nb\text{-}nc}$) and handoff call probability ($P_{hb\text{-}nc}$) are given by,

$$p_{nb-nc}=\frac{\sum_{n_2=0}^{C-K}\frac{\rho_n{}^K}{K!}*\frac{\rho_h^{n_2}}{n_2!}+\sum_{n_1=0}^{K-1}\frac{\rho_n{}^{n_1}}{n_1!}*\frac{\rho_h^{C-n_1}}{(C-n_1)!}}{\sum_{n_1=0}^{K}\frac{\rho_n{}^{n_1}}{n_1!}\sum_{n_2=0}^{C-n_1}\frac{\rho_h^{n_2}}{n_2!}} \quad (1)$$

$$p_{hb-nc}=\frac{\sum_{n_1=0}^{K}\frac{\rho_n{}^{n_1}}{n_1!}*\frac{\rho_h^{C-n_1}}{(C-n_1)!}}{\sum_{n_1=0}^{K}\frac{\rho_n{}^{n_1}}{n_1!}\sum_{n_2=0}^{C-n_1}\frac{\rho_h^{n_2}}{n_2!}} \quad (2)$$

1.1.2. Cutoff Priority Scheme. In this scheme, instead of putting limitation on the number of new calls, we base on the number of total on-going calls in the cell to make a decision whether a new arriving call is or is not accepted. The scheme works as follows.

Let *m* denote the threshold, upon a new call arrival. If the total number of busy channels is less than *m*, the new call is accepted; otherwise, the new call is blocked. The handoff calls are always accepted unless no channel is available upon the arrivals. Then, the expressions for new calls blocking probability and handoff dropping probabilities are,

$$p_{nb-cop}=\frac{\sum_{j=m}^{C}\frac{(\rho_n+\rho_h)^m\rho_h^{j-m}}{j!}}{\sum_{j=0}^{m}\frac{(\rho_n+\rho_h)^j}{j!}+\sum_{j=m+1}^{C}\frac{(\rho_n+\rho_h)^m\rho_h^{j-m}}{j!}} \quad (3)$$

$$p_{hb-cop}=\frac{\frac{(\rho_n+\rho_h)^m\rho_h^{C-m}}{C!}}{\sum_{j=0}^{m}\frac{(\rho_n+\rho_h)^j}{j!}+\sum_{j=m+1}^{C}\frac{(\rho_n+\rho_h)^m\rho_h^{j-m}}{j!}} \quad (4)$$

1.1.3. Number of flows at the wireless Access Router. Now, we can use expressions (1) to (4), given for P_{nb} and P_{hb} for both CAC models, to obtain the average number of new calls (N_n) and handoff calls (N_h) for each case. Therefore, we use the expression for the average number of elements over an Erlang Loss Queueing system given by (5),

$$N=\frac{\lambda(1-P_b)}{\mu} \quad (5)$$

Then, we obtain N_n and N_h from P_{nb} and P_{hb} respectively for each CAC model as,

$$N_n=\frac{\lambda_n(1-P_{nb})}{\mu_n} \quad (6)$$

and,

$$N_h=\frac{\lambda_h(1-P_{hb})}{\mu_h} \quad (7)$$

Then, the mean total number of users on the cell for each CAC model is calculated as,

$$N_t=N_n+N_h \quad (8)$$

1.2. Mean Packet Delay for the classification process

We have obtained a stochastic model for the classifier mean packet delay, which is based on the states transition diagram (STD) of the process. Also, to evaluate our approach, we have developed classification models for two different QoS support approaches. These are the Standard Integrated Services [5] and a new approach named IntServ6 proposed in [6]. Some key differences between IntServ and IntServ6 are that in IntServ6 the five-tuple reading and its calculation are not performed at the router; meanwhile in IntServ these actions are required. In addition, the hash table searching in IntServ6 uses a tagging scheme; while this process in IntServ uses a matching scheme (see [7]). Thus, the States Transition Diagram (STD) changes for each technology (IntServ and IntServ6); therefore, the obtained mathematical expressions for each one also change. STDs can not be shown by paper space requirements. Time parameters for the obtained expressions are: t_q (five-tuple reading time), t_h (hash number calculation time), t_{comp} (five-tuple comparison time), t_{comph} (comparison between hash number and a hash table row on IntServ), t_e (packet routing time) and t_i (hash table verification time).

Thus, the obtained expression for the mean service time on IntServ is,

$$\bar{x}_{co} = t_q + t_h + mt_{comph} + t_e + t_{comp} m C_h^2 \tag{9}$$

And the expression for service time on IntServ6 is,

$$\bar{x}_c = t_i + t_e + t_{comp} m C_h^2 \tag{10}$$

Equations (9) and (10) have a C_h parameter, which represents the collision rate. This parameter is introduced due to the use of a hash number to identify the flows on both technologies: IntServ and IntServ6. The collision rate is defined in [8] and is calculated as,

$$C_h = 1 - \frac{N * \left(1 - \left(\frac{N-1}{N}\right)^m\right)}{m} \tag{11}$$

Where N is the hash table size and m is the number of flows.

For simplicity, we suppose that packets arrive in bursts and, at the same time, these burst arrive with a Poisson distribution. Also, we suppose that our router has a memory cache that holds each burst. It allows performing only one classification process per each packet burst. Thus, we can assume that the arrivals to the classification process have a Poisson distribution. Then, we can model the classification process as an M/G/1 tail. We have considered a general distribution for the service discipline because there is a geometric distribution on the collision resolution process for both, IntServ and IntServ6, and also a geometric distribution in the hash table searching on IntServ.

However, as we want to use queuing networks principles to build our router model in future studies; we should obtain an M/M/1 model. Therefore, it is necessary to do some assumptions. Normally, the classifier looks for a hash number on the collision table if a collision has occurred. Time search for this case has a geometric distribution. But, if we consider that flow reservations are randomly created or dropped, we can assume that the flow allocation on the collision resolution table of IntServ and IntServ6 is randomly and exponentially distributed. We can make similar assumptions for the hash table of IntServ. Thus, we can assume that the classifier service time has a Poisson distribution. Then, under these assumptions, we can assume an M/M/1 tail for a specific session classification process. To perform this model, we can assume the same mean service time obtained for the M/G/1 model because the mean packet delay of the classification process is the same for all packets independently from the flow it belongs to (see equations (9) and (10)).

1.3. Mean packet delay on the scheduler

A significant volume of work in the literature has been concerned with evaluating the deterministic worst-case delay guarantees that *Generalized Processor Sharing* (*GPS*) algorithms can provide when the burstiness of the traffic feeding them is bounded (mostly shaped by a leaky bucket). However, little work has been performed on analyzing the stochastic delay bounds of such packetized policies under a general probabilistic traffic model. This has been mainly due to the difficulty of stochastic modeling of the complex behavior of a *Fair Queueing* (*FQ*) algorithm. The worst case delay bound was calculated by Parekh and Gallager. However, other studies have been looking for models of statistical bounds for packet delay on *GPS* scheduler based networks. These approaches [9, 10] have been based on the *Exponentially Bounded Burstiness* theory (*EBB*) developed by Yaron and Sidi [11], which assume arrivals of exponentially distributed bursts and consider that when the inputs to an isolated network element are all bounded, they result in bounded outputs, assuring exponentially decaying distributions delays. Finally, another work was developed by Zhang et al. in [12]; they obtained a stochastic model for upper delay bound of *PGPS* algorithms; that work was based on the notion of *feasible partition*. This model has packet arrivals with a Poisson distribution and packet lengths with a General Distribution. Zhang's approach obtained a model for a specific *i* session of a *PGPS* scheduler as M/G/1 tail and obtained expressions for mean time delay, variance and mean delay upper bound. This model was also particularized for an M/M/1 tail with the same parameters mentioned before. We have used this M/M/1 tail for our modeling and the expression used for the upper delay bound is,

$$\bar{d}_i \le \frac{L_i}{w_i C - \lambda_i L_i} \tag{12}$$

Where $\bar{d}_i$ is the mean packet delay for the i session, $\bar{L}_i$ is the average packet length for session i, C is the total output server capacity (in bits per second), w_i is the flow weight and λ_i is the arrival average rate for the i session.

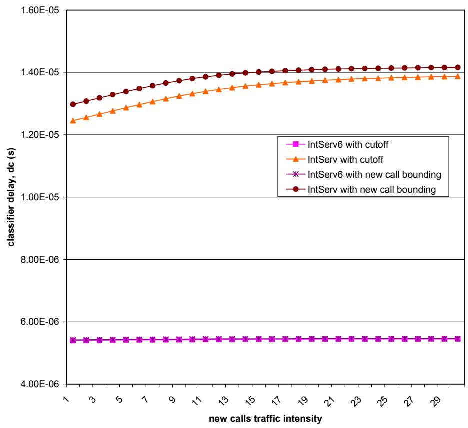

Figure 2. Mean packet delay for IntServ and IntServ6 under two CAC schemes vs. new calls traffic intensity.

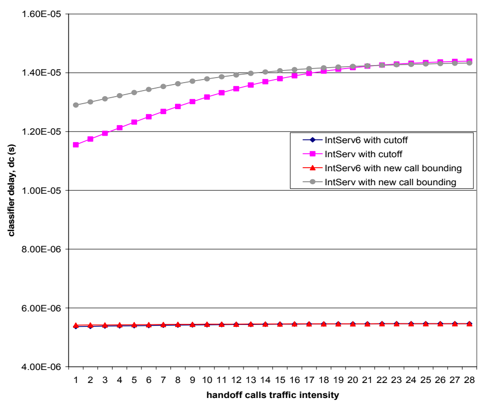

Figure 3. Mean packet delay for IntServ and IntServ6 under two CAC schemes vs. handoff traffic intensity.

2. Results and conclusions

In this paper we modeled the mobility effect over the mean packet delay bound at a network layer level. Thus, we have obtained a more complex model that the model obtained in [4]. For evaluation purposes, we have simulated the mean packet delay until the classification process because the scheduling process is the same for both QoS approaches. Thus, it does not represent a difference between both proposals. Meanwhile, the classification process is the main difference between IntServ and IntServ6. Then, it is more interesting to show the classification behavior on the router. As a result, Fig. 2 and 3 show the simulation results for the mean packet delay behavior with respect to the new calls and handoff calls traffic intensity for two CAC schemes and both QoS approaches (IntServ and IntServ6). For simulations we used an IXP 1200 network processor parameters and a typical C of 30 resources per cell. We can observe that different CAC schemes affect in different way the mean packet delay on the router. Thus, when the new calls traffic increases, the classification delay increases in a similar way for both CAC schemes. On the other hand, when the handoff calls traffic increases, the difference between the mean packet delay parameter in both CAC schemes is higher than the new calls traffic variation case. Also, the QoS approach used in the network is important to obtain a good network response to several mobility issues. Thus, IntServ6 gives a more independent response with respect to mobility parameters variation. On the other hand, IntServ has a high dependence with mobility issues such as CAC schemes and new calls and handoff calls traffic intensity variation.

3. References

[1] Y. Fang, I. Chlamtac, and Y. Lin, "Teletraffic analysis and mobility modeling for PCS networks", *IEEE Transactions on Communications*, vol. 47, pp. 1062-1072, 1999.

[2] Y. Fang, I. Chlamtac, and Y. Lin, "Channel occupancy times and handoff rate for mobile computing and PCS networks", *IEEE transactions on Comput.*, vol. 47, pp. 679-692, 1998.

[3] Y. Fang and Y. Zhang, "Call admission control schemes and performance analysis in wireless mobile networks", *Vehicular Technology, IEEE Transactions on*, vol. 51, pp. 371-382, 2002.

[4] F. Yu, L. Ma, and V. Leung, "Effects of Mobility on Packet Level Performance of Connection-Oriented Multimedia Traffic over Packet-Switched Cellular Networks", *IEEE WCNC*, pp. 2171-2176, 2004.

[5] R. Braden, "Integrated Services in the Internet Architecture: an Overview", IETF, RFC 1633 June 1994.

[6] J. Padilla, M. Huerta, J. Paradells, and X. Hesselbach, "IntServ6: An Approach to Support QoS over IPv6 Networks", *Computers and Communications, 2005. ISCC 2005. Proceedings. 10th IEEE Symposium on*, pp. 77-82, 2005.

[7] J. Kaur, *Scalable Network architectures for providing per-flow service guarantees*. Austin, Texas: Dissertation. The University of Texas at Austin, 2002.

[8] Z. Wang, *Internet QoS: Architectures and Mechanisms for Quality of Service*, 1st edition ed: Morgan Kaufmann, March 15, 2001.

[9] O. Yaron and M. Sidi, "Generalized Processor Sharing Networks with Exponentially Bounded Burstiness Arrivals", *High Speed Network*, pp. 628-634, 1994.

[10] Z. Zhang, D. Towsley, and J. Kurose, "Statistical Analysis of Generalized Processor Sharing Sheduling Discipline", presented at SIGCOMM'94, London, England, 1994.

[11] O. Yaron and M. Sidi, "Performance and Stability of Communication Networks via Robust Exponential Bounds", *IEEE/ACM Transactions on Networking*, vol. 1, pp. 1993, 1993.

[12] J. Zhang, T. Hamalainen, and J. Joutsensalo, "Stochastic Analysis of Upper Delay Bound of GPS-Based Packetized Fair Queueing Algorithms", presented at ICON 2004, 2004.