A Precision Predictive Call Admission Control in Packet Switched Multi-Service Wireless Cellular Networks

Robert G. Fry and Abbas Jamalipour
School of Electrical & Information Engineering
The University of Sydney, NSW 2006, Australia
{rfry,abbas}@ee.usyd.edu.au

***Abstract*— The increasing variety and complexity of traffic in today's mobile wireless networks means that there are more restrictions placed on a network in order to guarantee the individual requirements of the different traffic types and users. Call Admission Control (CAC) plays a vital role in achieving this.**

In this paper we propose a Precision Predictive Call Admission Control (PPCAC) scheme in which existing methods to determine the call occupancy distribution in each cell are modified in order to predict the number of calls of each individual service that are underway. Information about the channel usage of each service is then used to calculate the capacity of the cell which will ensure that each services' bit error rate (BER) requirements are met. Priorities assigned to each service are used to allocate the network capacity. An expression for the handoff dropping probability is derived and the acceptance rate for each service is calculated in order to guarantee that the required Quality of Service (QoS) is met. Each call is then accepted with equal probability throughout the duration of a control period. The new call acceptance rate is adjusted to ensure that the dropping probability remains fixed. The complexity of this algorithm is offset by the fact that most of the computations can be completed offline and it only requires readily available network measurements.

Simulations conducted in a CDMA environment with voice, short message service (SMS) and world wide web (WWW) sources show that the proposed CAC can successfully meet the hard restraint on the dropping probability and guarantee the required BER for multiple services.

I. Introduction

One of the most challenging issues in current and next generation wireless cellular networks is the ability to guarantee the QoS. This has to be achieved in an environment that is extremely variable in terms of the physical medium, the behaviour of the users and the actual traffic. It also has to efficiently use the available resource in order to be economically viable. One of the ways that this is achieved is through the use of a CAC algorithm, which restricts access to the network in such a way as to guarantee the QoS of existing calls whilst maximizing the network utilization. With the ever-increasing volume and diversity of today's wireless cellular traffic, CAC schemes need to be dynamic, fast and able to handle a variety of traffic loads and types. A highly mobile and varying traffic load dramatically increases the amount of handoffs that can be expected as well as the number of new calls trying to access the network. This variability makes the design of an effective CAC algorithm extremely difficult.

Wu, Wong and Li [1] propose a method that guarantees the handoff dropping probability for voice calls in a circuit switched network. Dropping probability is defined as the probability that a call is terminated prematurely when a handoff is attempted. The problem becomes considerably more complex when a multi-service network is considered. This paper deals with developing their technique in such a way that the QoS of a variety of packet switched services, such as those proposed in [2], can be simultaneously guaranteed whilst providing a high level of channel utilization.

This work was supported by the Australian Research Council and SingTel Optus Pty Limited

In circuit-switched networks, calls are allocated a maximum amount of bandwidth for the duration of each call. When traffic of a variable nature, such as packet switched traffic is introduced, this static bandwidth allocation is no longer suitable as the limited wireless capacity is under-utilized. In order to calculate the maximum number of calls that the network can service at any particular time alternate methods are required. Using the procedure in [1] we can model the number of users of each service active at any given time. The number of channels that this corresponds to can then be determined using the channel activity models for each service. This describes the channel usage of the service by a matrix of probabilities.

Since there are multiple services, the number of channels that can be allocated to a service is no longer just a matter of how much bandwidth is available. Dividing the total capacity up in a certain ratio is not a viable option as the limited radio channels would be severely under-utilized given the highly diversified traffic load expected in current and future networks.

A number of methods exist that allocate bandwidth to services or users dynamically [3–5] but fail in some way to meet all the required objectives of a next generation multi-service wireless cellular network.

The algorithm proposed by Chen, Kumar and Kuo [3] determines the resource reservation threshold based on neighbouring cell higher-priority calls which are likely to handoff. Although this ensures that higher priority calls achieve a better handoff dropping probability, it cannot guarantee a fixed level for increasing call arrival rates.

Lim, Cao and Das [4] propose a method that reserves bandwidth in cells depending on which direction the call is predicted to be travelling. This can guarantee the handoff dropping probability but sacrifices channel utilization as bandwidth is reserved in cells that the call will not necessarily handoff to. Also only one level of handoff dropping probability can be fixed for all services. In other words, it can not guarantee different QoS levels for multiple services.

The algorithm presented by Cao and Li [5] does provide individual QoS guarantees but it assumes that the user always generates enough traffic to fully consume the allocated bandwidth. This assumption is not suitable for services with traffic that generates an ON, OFF pattern, as it wastes valuable bandwidth during the OFF periods. For example, with a voice or web browsing call, data is only transmitted when the user is either talking or downloading a web page. For the rest of the time, while the user is either listening or reading, the channel is idle. This assumption makes any statistical multiplexing of channels impossible.

Although the objectives for all of the aforementioned papers are essentially the same, the environment and traffic for which they are proposed is not. The objectives of PPCAC is to guarantee the handoff dropping probability and BER's for multiple services, whilst maximizing the new call acceptance rates and channel utilization in a realistic multi-service current-to-next-generation mobile wireless cellular environment. We define the dropping probability to be the probability that a call is dropped as it attempts to handoff to a neighbouring cell or exceeds its required BER.

PPCAC uses a proactive approach to CAC which ensures that new calls are spread uniformly throughout the control period and can guarantee the dropping probability regardless of traffic load. Since there is no capacity reserved for handoffs efficient use is made of the available resources. The differing QoS requirements for each service can also have an effect on the number of channels that can be active at one time. For example, the BER experienced by a particular channel is proportional to the number of active channels, so that for a certain BER requirement the total capacity of a cell might be reduced.

In this paper, we expand on the algorithm proposed in [1] and adapt it for use in a multi-service environment. To do this we employ the original algorithm in parallel, each one operating on a different service. These generate the occupancy distributions for each service. From these we determine the acceptance rates (which is the rate with which new calls are accepted) of each service that will guarantee the dropping probability for next control period. At the end of each control period readily available network statistics are used to update the algorithm parameters. The required statistics are the number of calls of each service that are underway at the beginning of each control period.

The time dependant dropping probability in each cell is estimated using the call transition probabilities between cells up to three hops away. A hop is when a call moves between adjacent cells. We limit the number of hops to three as the likelihood of calls moving much further in a single control period is negligible. A closely packed hexagonal cell structure is considered. Using the QoS requirements, expressions for the dropping probabilities are developed and used to admit new calls to the network uniformly over the following control period.

From [1] the time dependant mean and variance of the occupancy distribution is obtained. This translates to a Gaussian distribution for large system sizes, however the limited capacity of cells means that this distribution is distorted, especially when the network is operating at close to or at full capacity. For a single service, circuit switched network the modified distribution can be easily determined, as the maximum capacity is known. In a multi-service network where capacity is allocated to each service dynamically, the 'hard' limit is not so readily obtainable. This is made more difficult when calls are packet switched and generate bursty traffic so that at any particular moment in time a single call could be using any number of channels.

Using channel activity models for each service, which describe the ON/OFF characteristics of the traffic, the maximum possible capacity can be allocated to each service without violating the QoS bounds. Priorities are assigned to each service to ensure that this is achieved without undue bias to a particular service.

The rest of the paper is organized as follows. In Section II, we describe the algorithm and traffic models. Some numerical results are presented in Section III. Conclusions and further works are detailed in Section IV.

II. Method

PPCAC uses predicted cell traffic to determine the new call acceptance rate for each service that guarantees the required QoS. Cell traffic is predicted using periodic network measurements and simple learning algorithms. We assume packet switched, Poisson distributed new calls of service type i arrive at a rate of λ_k^i in cell k. Call length is exponentially distributed with average $1/\mu_k^i$ except for WWW traffic which has a geometrically distributed call length. h_{jk}^i is the handoff rate of service i from cell k to cell j. The length of the control period is T seconds and P_{QoS}^i is the threshold for the handoff dropping probability of service i.

The derivation for the single-call of service type i transitional probabilities $f_{jk}^i(t)$ is detailed in [1] and is obtained by summing the path contribution q over all possible paths from cell k to cell j that do not exceed three hops:

$$q_b^i(t) = \frac{1}{b!}\left(\frac{h_{jk}^i t}{6}\right)^b \exp\left[-\left(h_{jk}^i + \mu_k^i\right)t\right] \tag{1}$$

where b is the number of hops between cell k and cell j. q is the fraction of calls of a particular service type i from cell k that can be expected in cell j after t seconds. Equation (1) is only valid for uniform traffic loading which we assume is the case for each service i. For the non-uniform traffic loading case a more complex analysis is required.

These probabilities are used to model the movement of users about the network. For simplicity the effect of soft handoffs is not considered and will be the subject of future work. The number of calls of a particular service type underway in each cell can be described by the distribution developed in [1]. The mean and variance of the call occupancies are derived by considering the number of calls underway in cell k for each service type i at the beginning of the control period, n_{k0}^i, which are obtained by network measurements, the average number of ongoing local calls, $n_{k0}^i f_{kk}^i(t)$, and the average number of new local calls, $a_k^i g_{kk}^i(t)/T$ in each cell during the control period. $g_{kk}^i(t) = \int_0^t f_{kk}(t-t')dt'$ which is the integrated transitional probability. We assume that the average number of handout and hand-in calls balance each other out so the net result is negligible. Even with mobility equilibrium the occupancy distribution is not stationary because of the stochastic nature of call arrival and termination rates, this results in the distribution being a function of time. The mean is given as [1]:

$$n_k^i(t) = f_{kk}^i(t)n_{k0}^i + g_{kk}^i(t)a_k^i/T + \frac{1 - f_{kk}^i(t) - \mu_k^i g_{kk}^i(t)}{1 - f_{kk}^i(T) - \mu_k^i g_{kk}^i(T)} n_{kb}^i(T) \tag{2}$$

and the variance as:

$$\sigma_k^i(t)^2 = n_k^i(t) - f_{kk}^i(t)^2 n_{k0}^i \tag{3}$$

where $n_{kb}^i(T)$ is the number of background calls, i.e. calls that handoff from other cells, during a control period, into cell k and $\frac{1-f_{kk}^i(t)-\mu_k^i g_{kk}^i(t)}{1-f_{kk}^i(T)-\mu_k^i g_{kk}^i(T)}$ is the fraction of $n_{kb}^i(T)$ that can be expected in cell k t seconds after the beginning of each control period.

In order to develop an expression for the dropping probability of a call from the occupancy distributions, the maximum capacity for each service is required. In CDMA networks, this capacity is not a fixed value as it depends on how many calls of other service types are underway at the same time, i.e. it is only limited by interference levels unlike FDMA and TDMA networks which are limited primarily by bandwidth [6]. The capacity of a cell is the number of connections that can simultaneously exist with an acceptable level of mutual interference. This is commonly referred to as the soft capacity. Different service types contribute different levels of interference so the capacity is dependant on the type of calls as well as the number of calls.

One of the contributions of this work is the method we use to map the number of calls to the bandwidth usage in each cell. In achieving this we can then calculate the dropping probabilities for each service. Randhawa and Hardy [7] propose a method for determining the maximum number of channels that can simultaneously be ON whilst not causing the BER of the channel to exceed a certain limit. The number of connections that this corresponds to can be determined using the channel activity models of that service. These models describe the length of time that a call of a given service will occupy a given number of channels.

The service types used in this paper are voice, SMS and WWW traffic. The voice source only transmits when the user is talking and uses silence detection to determine when the channel is idle. During the talk spurt data is transmitted at a peak rate of 12.2 kbps. The length of the talk spurt is exponentially distributed with a mean of 3 seconds and the OFF time is exponentially distributed with a mean of 4.5 seconds. This corresponds to a channel activity factor of 0.4. Each SMS message is transmitted in its entirety at a constant bit rate. Message size is modelled by a geometric distribution with a mean of 280 bytes [8]. WWW traffic alternates between a download state and a reading state. The download state consists of periods of activity where a geometrically distributed number of packets, with a mean of 25, of a Pareto distributed size with mean 480 bytes is broken by periods of inactivity which are exponentially distributed with a mean of 0.0277 seconds. The time spent in the reading state is exponentially distributed with a mean of 12 seconds [8].

In order to determine the dropping probability we need to be able to determine the capacity of the network. In a CDMA environment this capacity is a function not only of the available bit energy to noise density ratio, spectrum and bitrate, but also of the amount of data being transmitted at a certain point in time. Given that we know or can predict the number of users in the network we can also determine how much traffic is being transmitted using the channel activity factors of each service which give a reflection of the amount of time a channel is used to transmit during the call. We consider the uplink capacity as the limiting case. As the number of users transmitting at the same time increases the BER deteriorates. In order to prevent this from occurring we guarantee a maximum BER that can only be exceeded by a probability ϵ. This puts an upper bound on the number of channels that can be active at one time.

Using these models and the soft capacity as a function of the service BER (we consider in this paper a WCDMA network), the maximum number of connections for each service can be determined. Each service is allocated a priority w_i that corresponds to the percentage of the capacity it is assigned relative to the other services.

$$\sum_{i=0}^{I} w_i = 1 \tag{4}$$

where I equals the total number of services.

For each possible combination of connections for each service, the probability that the BER will be exceeded can be calculated. The largest set of connections is that which results in the BER being exceeded within the fixed tolerance level, ϵ. The BER experienced by a particular channel, $\mathrm{BER}(c)$ is a function of the number of other active channels c in the cell at that particular time. The maximum number of channels that can be ON simultaneously whilst satisfying the BER limit, BER_i is [7]:

$$C^i_{\max} = \max\{c \, : \, \mathrm{BER}(c) \le \mathrm{BER}_i\} \tag{5}$$

where

$$\mathrm{BER}(c) = Q\left\{\frac{\frac{W/B}{SNR} - \frac{\eta_t}{S} - (c-1) - E}{\sqrt{\mathrm{Var}}}\right\} \tag{6}$$

and W, B, SNR, η_t and S are the available spectrum, bitrate, signal to noise ratio, thermal noise and signal power respectively and $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$. Perfect power control is assumed for the uplink so that the equal power is received at the base station from each user. E and Var are the mean and variance of the outer cell interference [6].

The channel activity models $P_i(j,c)$ describe the probability that given j connections c channels will be ON. The probability that all $C^i_{\max}$ channels will be ON is likely to be small so that the maximum number of connections j_i of each service can be determined by satisfying

$$\sum_{c=0}^{C^i_{\max}} P_c > 1 - \epsilon \tag{7}$$

where $P_c = \prod_{x=1}^{I} P(j_x, c_x)$ and the summation is performed over all possible combinations of c_x that satisfy $\sum_{x=1}^{I} c_x = c$.

However the number of connections required by each service at different times will be highly variable. This could result in the admission rate being set at a rate that is not the most resource efficient. Instead of fixing this connection limit we propose a scheme which uses the service priorities to vary the maximum connection levels according to the relative channel usage between services.

There are two different conditions of the network that we need to consider; when $\sum_{i=1}^{I} n_i(t) > N_T$ or $\sum_{i=1}^{I} n_i(t) <= N_T$. $n_i(t)$ is the predicted connection level for service i at time t and N_T is the maximum number of connections that can exist simultaneously. Hereafter the reference to t will be omitted as all calculations are evaluated at the same point in time.

Firstly, when $\sum_{i=1}^{I} n_i > N_T$. If this is the case then the network is likely to exceed its maximum connection levels and calls will be dropped. In a static priority scheme the expected maximum connection level for service i is $w_i N_T$ and for the case that $n_i > w_i N_T$ for all i this remains the same. For the case though that any services j, exist where the connection level $n_j < w_j N_T$ the minimum dropping probability is overestimated if the maximum connection level N_j is set to $w_j N_T$. For services i for which $n_i > w_i N_T$, N_i is determined by subtracting a proportion of the amount by which N_T is exceeded which corresponds to priority of the service i, w_i. This is shown below:

$$N_i = n_i - \frac{w_i}{\sum\limits_{x \neq j} w_x}\left\{\sum_{x=1}^{I}(n_x - w_x N_T) + \sum_j w_j(N_T w_j - n_j)\right\} \tag{8}$$

where i corresponds to services for which $n_i > w_i N_T$ and j corresponds to services for which $n_j < w_j N_T$. If any service's utilization of the network $n_i = w_i N_T$ then N_i is set to n_i. For those services j, N_j is determined by adding a proportion, equivalent to the priority w_j, of the amount that n_j is less than $w_j N_T$, thus,

$$N_i = n_i + w_i(N_T w_i - n_i) \tag{9}$$

where i corresponds to services for which $n_i < w_i N_T$.

Lastly, when $\sum_{i=1}^{I} n_i <= N_T$. Each services' maximum connection level is determined by increasing the usage level by a proportion, dictated by the priority w_i, of the available free bandwidth.

$$N_i = n_i + w_i(N_T - \sum_{x=1}^{I} n_x) \tag{10}$$

This scheme ensures that the bandwidth utilization achieved is a maximum and the dropping probability is reduced to its minimum possible level.

Once this has been determined, the dropping probability can then be estimated by considering the modified occupancy distribution at the region of full occupancy, i.e. when $x = n/N = 1$. This is given in [1] as,

$$D_i = \frac{P_i(x)}{N_i} \tag{11}$$

where $P_i(x)$ is the modified occupancy distribution, n is the number of calls underway and N_i is the maximum number of channels available in the cell. $P(x)$ is a function of the time dependant mean and variance of the channel occupancy given in [1] which is also a function of the acceptance rate. Using the approach outlined above we can perform a simple mapping procedure that results in the occupancy distribution describing the number of channels as opposed to the number of connections underway in a cell at a certain time.

However, since the value of n_k^i (n_k^i represents the number of connections of service type i in cell k), at full occupancy is not a constant or known with certainty the expression (1) is no longer accurate. As shown previously, the derived maximum connection numbers can be exceeded with probability less than or equal to the fixed tolerance level, ϵ. Therefore, the modified dropping probability is more accurately described by,

$$D_k^i = \frac{P_k^i(1)}{(1+\epsilon_i)N_k^i} \tag{12}$$

where $P_k^i(1)$ is the modified distribution of connections evaluated at the point of full occupancy and N_k^i is the maximum number of connections available in cell k for a particular service i.

The acceptance rate a_k^i/T is determined by solving $\widetilde{D}_k^i = P_{\text{QoS}}^i$, where $\widetilde{D}_k^i$ is the average dropping probability over a control period. Obviously a_k^i cannot be negative so if the resulting solution is less than zero, a_k^i is set to zero. If a_k^i/T is greater than the arrival rate of service i in cell k i.e. $a_k^i/T > \lambda_k^i$ then a_k^i is set to $\lambda_k^i T$.

III. Results

For each control period the acceptance rate for each service that results in the required dropping probability is calculated. Cell occupancies of each service type n_{k0}^i are measured at the beginning of each period. Using the cell occupancies, $n_k^i(T)$ can be estimated and updated each control period via

$$\begin{aligned} n_k b^i(T) \leftarrow & (1-r) n_k b^i(T) \\ & + r\left\{ n_k b^i(T) - f_{kk}^i(T) n_{k0}^i - g_{kk}^i(T)\frac{a_k^i}{T}\right\} \end{aligned} \tag{13}$$

where r is the learning rate of the algorithm.

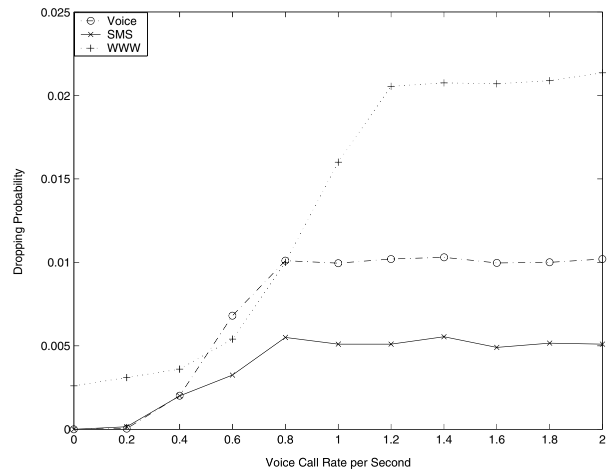

Fig. 1. Dropping probabilities for different services in PPCAC.

The secant adaption of the Newton-Raphson method [9] is used to calculate the new acceptance rate for the next control period.

The following parameters were used in Figs. 1-3:

- The call termination rates for each service, $\mu^{\text{voice}} = 0.05s^{-1}$, $\mu^{\text{WWW}} = 0.01s^{-1}$, $\mu^{\text{SMS}} = 0.2s^{-1}$.
- The call handoff rates for each service, $h^{\text{voice}} = 0.01s^{-1}$, $h^{\text{WWW}} = 0.005s^{-1}$, $h^{\text{SMS}} = 0.001s^{-1}$.
- The control period, $T = 20$s.
- The QoS thresholds for each service, $P_{\text{QoS}}^{\text{voice}} = 0.01$, $P_{\text{QoS}}^{\text{WWW}} = 0.02$, $P_{\text{QoS}}^{\text{SMS}} = 0.005$. These are the maximum tolerable dropping probabilities.
- The priority weighting values for voice, WWW and SMS sources are: 0.5, 0.3, and 0.2, with BER limits: 10^{-2}, 10^{-5}, and 10^{-5} [2] respectively.
- The channel usage probabilities P_x^i for each service i and number of channels x that are ON are; $P_0^{\text{voice}} = 0.575$, $P_1^{\text{voice}} = 0.425$, $P_0^{\text{WWW}} = 0.322$, $P_1^{\text{WWW}} = 0.649$, $P_1^{\text{SMS}} = 1$ [8].

Simulations were performed by averaging over six samples each with a simulation time of 1 hour. The number of background calls in the previous period is updated using an exponential smoothing technique [1]. Transitional probabilities are then calculated.

Fig. 1 shows the dropping probability of the different services over an increasing new voice call arrival rate in PPCAC. The new call arrival rates of the WWW and SMS sources are both constant with two calls per second. PPCAC maintains the dropping probability of each service at the specified QoS threshold with the required stability and accuracy, even for high call arrival rates. The dropping probability of WWW calls is slightly higher than the specified value which is a result of the learning rate of the algorithm and the mean length of the WWW calls. Once WWW calls are admitted to the network they are more likely to remain for a longer period of time in one cell than other services, if the algorithm is too generous initially overshoot will occur. Modifying the learning rate will ensure this does not occur but will reduce the admission rate slightly and therefore decrease the capacity utilization. The overlap of the voice and WWW dropping probabilities that occurs around an arrival rate of 0.6 can be attributed to the fact that PPCAC adjusts its approach according to how much each particular

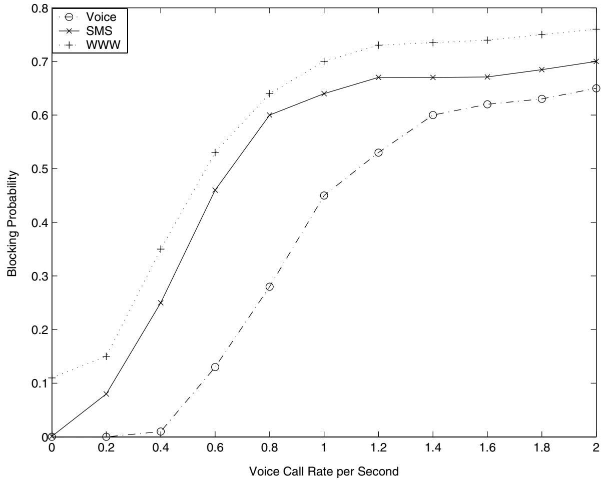

Fig. 2. Blocking probabilities for different services in PPCAC.

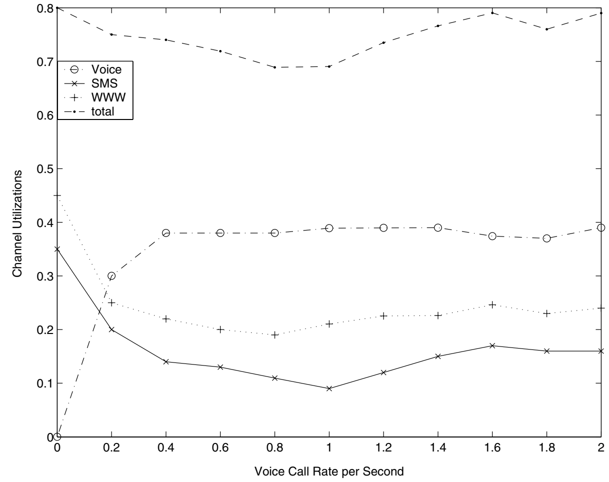

Fig. 3. Channel Utilizations for different services in PPCAC.

service is using the network at a given point in time. We expect that this anomaly will disappear over a longer simulation time. Using the information available at the beginning of each control period PPCAC can determine whether the arrival rates of each service are causing the network to approach or exceed the QoS limits that have been specified. If it is, the admission rate of calls is reduced sufficiently to avoid the QoS requirements not being met.

Fig. 2 shows the new call blocking probabilities of PPCAC, given by:

$$P_b^i = 1 - \frac{a^i}{\lambda^i T} \tag{14}$$

for each service as a function of the new voice call arrival rate. The results here are comparable to those from existing schemes [1, 3–5, 7]. As the amount of traffic requesting access to the network increases, PPCAC maintains a control on the dropping probabilities at the expense of the new call blocking probabilities. The degradation of the blocking probabilities is acceptable as it ensures that every call the network accepts can be guaranteed its QoS level whilst ensuring that the number of new calls accepted is as high as possible. This guarantee does not extend to calls that are dropped for other reasons such as loss of signal power in a network 'blind spot' or other physical layer limitations.

Fig. 3 shows that the proposed scheme maintains a high level of channel utilization over a wide range of voice call arrival rates. The total is the summation of all services. The priority weighting scheme effectively controls the proportion of the channel that each service receives. The lower priority calls sacrifice bandwidth to allow for the higher priority calls without greatly effecting the overall channel utilization. The proposed scheme remains stable over the entire range of arrival rates.

IV. Conclusion

In this paper we present a PPCAC for packet switched multi-service wireless cellular networks that can individually guarantee the QoS of each service in terms of the dropping probabilities and BER's whilst maximizing the channel utilization through the use of statistical multiplexing and maintenance of a competitive new call blocking rate. We improve the utilization using a priority-weighting scheme that allows the relative capacity of the network that each service receives to be controlled. Unlike other schemes, we use a proactive approach in making the admission decision, and set it to the value we compute over the previous control period. This ensures that new calls are uniformly spread over the length of the control period which leads to a stable CAC scheme that performs extremely well regardless of network load.

The selection of control period duration is critical to ensuring that the best performance is achieved. Its value is closely related to the mean duration of the calls of each service and the overall network load. It can be selected by testing the algorithm with a range of values which increases the computational load and future work is required to perhaps implement an adaptive control period length.

We are currently developing traffic models to coincide with the proposed UMTS standards [2] and conducting more comprehensive testing under a variety of traffic and channel conditions, including evolving handoff rates.

References

[1] S. Wu, K. Y. M. Wong, and B. Li, "A dynamic call admission policy with precision QoS guarantee using stochastic control for mobile wireless networks," *IEEE Transactions on Networking*, vol. 10, pp. 257–271, Apr. 2002.

[2] *ETSI TS 123 107 V4.3.0 (2002-01) Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture (3GPP TS 23.107 version 4.3.0 Release 4).* ETSI Technical Specification, 2002.

[3] H. Chen, S. Kumar, and C. C. J. Kuo, "Dynamic call admission control scheme for QoS priority handoff in multimedia cellular systems," in *Proc. IEEE WCNC'02*, pp. 114–118, Mar. 2002.

[4] S. Lim, G. Cao, and C. R. Das, "An admission control scheme for QoS-sensitive cellular networks," in *Proc. IEEE WCNC'02*, pp. 296–300, Mar. 2002.

[5] Y. Cao and V. O. K. Li, "Utility-oriented adaptive QoS and bandwidth allocation in wireless networks," in *Proc. IEEE ICC'02*, vol. 5, pp. 3071–3075, Apr. 2002.

[6] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, J. L. A. Weaver, and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Tech.*, vol. 40, pp. 303–312, May 1991.

[7] T. S. Randhawa and R. H. S. Hardy, "Performance evaluation of priority-based call admission control in multi-service DS-CDMA cellular networks supporting VBR video traffic," in *Proc. IEEE ICC'02*, vol. 5, pp. 3096–3105, Apr. 2002.

[8] A. Andreadis, R. Fantacci, G. Giambene, and F. Petiti, "An efficient uplink and downlink resource sharing scheme for wireless multimedia systems," *IEICE Trans. Commun.*, vol. E85-B, pp. 1512–1524, Aug. 2002.

[9] E. Kreyszig, *Advanced Engineering Mathematics*. John Wiley & Sons, 7th ed., 1993.