Utility-based Bandwidth Adaptation for Multi-Class Traffic QoS Provisioning in Wireless Networks

Ning Lu and John Bigham
Queen Mary, University of London

Abstract: Adaptive bandwidth allocation is becoming very attractive in wireless communications since it can dynamically adjust the allocated bandwidth of ongoing calls to cope with the network resource fluctuations. In this paper, we propose a utility-based bandwidth adaptation scheme for multi-class traffic QoS provisioning in wireless networks. With the proposed scheme, each call is assigned a utility function and depending on the network load the bandwidth of ongoing calls are upgraded or degraded so that the achieved utility of each individual cell is maximized. Simulation experiments are carried out to evaluate the performance of the proposed scheme. Results show that our adaptive scheme is effective in both increasing cell utility and reducing the call blocking and handoff dropping probabilities of wireless networks.

1. Introduction

Bandwidth adaptation is one of the most promising methods to provide QoS guarantees in wireless networks. In the traditional non-adaptive network environment, once a call is admitted its allocated bandwidth is fixed throughout its lifetime; when a new or handoff call requests a certain amount of bandwidth, the network rejects the call if there is not sufficient bandwidth available. However with adaptive bandwidth allocation, when a new or handoff call comes and the network is overloaded, the allocated bandwidth of ongoing calls can be degraded to smaller values, allowing the new or handoff call to be accepted [1]; this can reduce both call blocking and handoff dropping probabilities. On the other hand, when an ongoing call is terminated due to its completion or outgoing handoff, the released bandwidth can be utilized to upgrade other ongoing calls [2].

2. Utility-based Bandwidth Adaptation Modelling

In our wireless network model, the traffic offered to the network is assumed to belong to three classes:

- Class I – real-time non-adaptive multimedia traffic with stringent bandwidth requirements. Its utility function can be modelled by $u(b) = \begin{cases} 1, \text{ when } b \ge b_{min} \\ 0, \text{ when } b < b_{min} \end{cases}$, where b_{min} is the minimum bandwidth requirement.
- Class II – real-time adaptive multimedia traffic that can adjust their transmission rates to various network loads. Its utility function can be modelled by $u(b) = 1 - e^{-\frac{k_1 b^2}{k_2 + b}}$, where k_1 and k_2 are tunable parameters which determine the shape of the utility function and ensure that when the maximum requested bandwidth is received, $u \approx 1$.
- Class III – non real-time data traffic that is rather tolerant of delays. Its utility function can be modelled by $u(b) = 1 - e^{-\frac{kb}{b_{max}}}$, where k is a tunable parameter which determines the shape of the utility function and ensures that when the maximum requested bandwidth is received $u \approx 1$.

Each call is assigned to a utility function with shape depending on the characteristics of its traffic class. Bandwidth adaptation is performed based on each individual cell with fixed bandwidth capacity C; it can be decomposed into two processes – bandwidth degrades and bandwidth upgrades. Consider an overloaded cell with n ongoing calls, denote the i-th ongoing call's utility function as $u_i(b_i)$ $(1 \le i \le n)$ and its current allocated bandwidth as β_i. We define the i-th ongoing call's degradable utility function as $u_i^{\downarrow}(b_i) = u_i(\beta_i - b_i)$ $(0 \le b_i \le \beta_i)$ and its upgradable utility function as

$u_i^{\uparrow}(b_i) = u_i(\beta_i + b_i)$ $(0 \le b_i \le b_{i,\max} - \beta_i)$ where $b_{i,\max}$ is the maximum bandwidth requirement. We also apply an adaptation penalty function $p\_adapt(b_i)$ to represent the end-user dissatisfaction whenever a call is adapted; in our scheme, the adaptation penalty is a positive utility value proportional to the utility change due to bandwidth adaptation and is subtracted when calculating the total utility.

When a new or handoff call comes the allocated bandwidth of ongoing calls can be degraded to smaller values to accommodate the new or handoff call, thereby reducing the call blocking or handoff dropping probability. Denote the new or handoff call's utility function as $u_{n+1}(b_{n+1})$; the objective of bandwidth degrades is to degrade the bandwidth of the n ongoing calls and allocate bandwidth to the new or handoff call to maximize the sum of all calls' utilities subject to bandwidth constraints, namely

$$\text{maximize: } \left(\sum_{i=1}^{n}\left(u_i^{\downarrow}(b_i) - p\_adapt(b_i)\right)\right) + \left(u_{n+1}(b_{n+1}) - p\_adapt(b_{n+1})\right)$$

$$\text{subject to: } \left(\sum_{i=1}^{n}(\beta_i - b_i)\right) + b_{n+1} \le C \text{ and } 0 \le b_i \le \beta_i$$

When a call is terminated or handed off to another cell the released bandwidth can be utilized to upgrade the other $n-1$ ongoing calls such that the sum of their utilities is maximized, i.e.

$$\text{maximize: } \sum_{i=1}^{n-1}\left(u_i^{\uparrow}(b_i) - p\_adapt(b_i)\right) \qquad \text{subject to: } \sum_{i=1}^{n-1} b_i \le \beta \text{ and } 0 \le b_i \le b_{i,\max} - \beta_i$$

3. Bandwidth Adaptation Algorithm

The essence of bandwidth adaptation is to maximize the utility sum of a number of utility functions ($n+1$ utility functions for bandwidth degrades and $n-1$ utility functions for bandwidth upgrades). Without loss of generality, we propose an algorithm to maximize the total utility of n utility functions. Finding optimal solutions for such maximization problem is NP-hard and has exponential time complexity [3]. To support the real-time bandwidth adaptation in wireless networks our algorithm seeks to achieve near-optimal solutions with polynomial time complexity. We first quantize each utility function into line segments by dividing the bandwidth range into a fixed number of equal intervals. After quantization the utility function becomes a linear piece-wise function represented by a set of $<$bandwidth, utility$>$ points; thus the solution can be found in the non-dominated and non LP-dominated[1] points [4]. The utility maximization algorithm is described as follows:

Step 1: for $i := 1$ to n do

$u_i' := quantize(u_i)$ // quantize utility function u_i

$u_i'' := remove(u_i')$ // remove all dominated and LP-dominated utility points

Step 2: $u_{agg} := merge\_and\_sort(u_1'', \cdots, u_n'')$ // merge all utility functions' line segments into one // aggregated utility function and sort them by the // decreasing order of their slopes

Step 3: for $j := 1$ to $|u_{agg}|$ do // $|u_{agg}|$ is the number of line segments of u_{agg}

if $(b_{avail} \ge u_{agg}[j].bw)$ // if the available bandwidth is enough for line segment j

allocate $u_{agg}[j].bw$ to $u_{agg}[j]$ // allocate bandwidth to line segment j

else break;

Step 4: return $< b_i, u_i >$ // return the $<$bandwidth, utility$>$ profile for each utility function

[1] The concept of dominated and LP-dominated points can be found in [4].

4. Simulation Results

We use a 6×6 wrap-around cellular network model to evaluate the performance of the proposed scheme; each cell has a bandwidth capacity of 30 Mbps and the diameter of the cell is 1 km. New call arrivals are assumed to follow a Poisson distribution with mean rate λ and the call holding time is assumed to follow an exponential distribution with mean $1/\mu$. Mobile terminals (MT) can travel in one of six directions with equal probability and their speed is uniformly distributed between 10 and 60 miles per hour. Six representative groups of traffic are carefully chosen for the simulation; each traffic group is associated with an appropriate utility function and calls belonging to the same group are assumed to have the same utility function. The adaptation penalty is obtained by multiplying the bandwidth adaptation utility change by a fixed percentage which is 30% in our experiments. Table 1 shows the exact characteristics of the traffic used in the simulation. All six groups of traffic are generated with equal probability.

Table 1

Traffic characteristics for our simulation

Applic. Group	Traffic Class	Bandwidth Requirement	Average Connection Duration	Example	Utility Function (b is Mbps)
1	I	30 Kbps (CBR)	3 minutes	Voice Service & Audio-phone	$\begin{cases} 1, & b \geq 0.03 \\ 0, & b < 0.03 \end{cases}$
2	I	256 Kbps (CBR)	5 minutes	Video-phone & Video-conference	$\begin{cases} 1, & b \geq 0.25 \\ 0, & b < 0.25 \end{cases}$
3	II	1 - 6 Mbps (VBR)	10 minutes	Interact. Multimedia & Video on Demand	$1 - e^{-\frac{1.045b^2}{2.166+b}}$
4	III	5 - 20 Kbps (UBR)	30 seconds	E-mail, Paging & Fax	$1 - e^{-\frac{4.6b}{0.02}}$
5	III	64 - 512 Kbp (UBR)	3 minutes	Remote Login & Data on Demand	$1 - e^{-\frac{4.6b}{0.5}}$
6	III	1 - 10 Mbps (UBR)	2 minutes	File Transfer & Retrieval Service	$1 - e^{-\frac{4.6b}{10}}$

Our bandwidth adaptation scheme is compared with a rate-based borrowing scheme (RBBS) proposed in [1] and a non-adaptive scheme where a call must be allocated its maximum bandwidth to be admitted and once accepted its bandwidth cannot be changed throughout the lifetime.

Fig. 1 illustrates the utility comparison of the three schemes. The results show that our bandwidth adaptation scheme achieves more average cell utility[2] than other two schemes; the intuition behind this is that our scheme works in the fashion to maximize the sum of all calls' utilities in each individual cell.

Next, Fig. 2 shows the call blocking probability of the three schemes; it demonstrates how our scheme allows a significant improvement. At the arrival rate of 1.0 (calls/sec/cell), the blocking ratio is about 22% for our scheme, 28% for RBBS and 38% for the non-adaptive one; it has been decreased up to about 16% by introducing our adaptive scheme.

[2] Cell utility means the sum of all ongoing calls' utilities within a given cell.

Fig. 3 compares the handoff dropping probability. As can be seen, our scheme outperforms the non-adaptive one by reducing the handoff dropping probability up to about 7% and keeps it less than 7.5% as the call arrival rate increases to 1.0 (calls/sec/cell). RBBS achieves a lower dropping ratio than our scheme due to the fact that it reserves 5% bandwidth exclusively for handoff calls; however, this has decreased the bandwidth utilization.

To summarize, the overall performance of our bandwidth adaptation scheme is very attractive in both increasing cell utility and reducing call blocking and handoff dropping probabilities.

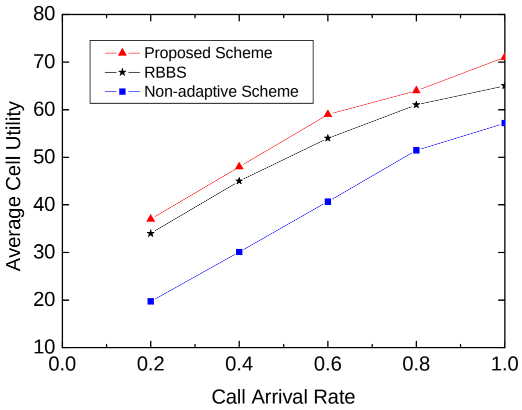

Fig. 1 Average cell utility

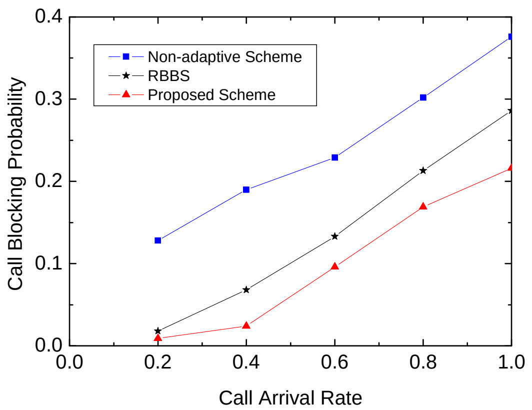

Fig. 2 Call blocking probability.

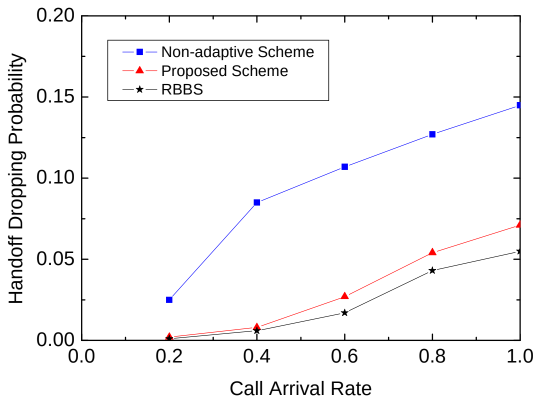

Fig. 3 Handoff dropping probability.

References:

[1] M. El-Kadi, S. Olariu, and H. Abdel-Wahab, "A Rate-based Borrowing Scheme for QoS Provisioning in Multimedia Wireless Networks", IEEE Transactions on Parallel and Distributed Systems, 13(2):156–167, Feb. 2002.

[2] K.-M. Ahn and S. Kim, "Optimal Bandwidth Allocation for Bandwidth Adaptation in Wireless Multimedia Networks", Computers & Operations Research, 30: 1917-1929, 2003.

[3] C. Lee, J. P. Lehoczky, R. Rajkumar and D. P. Siewiorek, "On Quality of Service Optimization with Discrete QoS Options", Proceedings of the IEEE Real-Time Technology and Applications Symposium, June 1999.

[4] H. Kellerer, U. Pferschy and D. Pisinger, "Knapsack Problems", Springer-Verlag, 317-322, 2004.