WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2272

RESEARCH ARTICLE

Admission control for WiMAX networks

Juliana Freitag Borin* and Nelson Luis Saldanha da Fonseca

Institute of Computing, University of Campinas, Campinas, 13089-971, SP, Brazil

ABSTRACT

This paper introduces admission control policies for WiMAX networks, which aim to reach three main goals: restrict the number of simultaneous connections in the system so that the resources available for the uplink scheduler are sufficient to guarantee the quality-of-service requirements for each connection, support the service provider expectations by maximizing the revenue, and maximize user satisfaction by granting additional resources. The optimal solution to meet these goals is non-deterministic polynomial-time hard and therefore cannot be solved in acceptable polynomial time. For this reason, both optimal and polynomial time heuristic solutions are introduced. Simulation experiments are used to evaluate and compare the proposed policies. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS
IEEE 802.16; quality of service; admission control

***Correspondence**
Juliana Freitag Borin, Institute of Computing, University of Campinas, Campinas, 13089-971, SP, Brazil.
E-mail: juliana@ic.unicamp.br

1. INTRODUCTION

Admission control mechanisms regulate the traffic load in the network so that the scheduler can furnish the bandwidth required by each connection. In IEEE 802.16 networks [1], often referred to as WiMAX networks, users request the establishment of a new connection by informing the desired type of service, along with a set of quality-of-service (QoS) requirements. Such service differentiation can result in users over-requesting resources to maximize their individual satisfaction. One effective way of encouraging users to choose the service that is most appropriate for their needs is through network pricing. Therefore, integrating pricing and admission control (AC) is an attractive alternative for service providers not only from an economic perspective but also from that of resources management.

This paper proposes various policies for AC in IEEE 802.16 networks, which were designed to maximize both service provider and user satisfaction. The proposed policies have the following goals: (1) to restrict the number of simultaneous uplink connections in order to avoid saturation of the uplink channel; (2) to meet service provider expectations by maximizing network revenue; and (3) to maximize user satisfaction (called utility in this paper) by providing resources beyond those required in the admission process.

The first AC policy presented in this paper is the simplest one, because it attains only the first goal. The decision strategy for the admission of new connections consists of guaranteeing that the sum of the minimum rate requirements of all accepted connections does not surpass the available capacity. This strategy guarantees that the resources available to the scheduler are sufficient to provide the bandwidth requirements for all accepted connections.

The second and third policies attain the first two goals. The second policy finds the optimal solution for the problem of admission of new connections and revenue maximization using integer linear programming (ILP), whereas the third uses a greedy approach to provide an easier-to-implement solution for the same problem.

The fourth and fifth policies attain all three goals. The fourth policy uses mixed ILP (MILP) formulation, and the fifth employs a heuristic algorithm to find solutions for the problem of admission of new connections and maximization of both revenue and utility.

Most AC solutions for IEEE 802.16 networks available in the literature [2–5] consider only minimum rate requirements in the decision process. The policies introduced in this paper differ from existing ones by the accountability of the overhead generated by the type of service being solicited, such as the overhead incurred by the bandwidth request mechanism specified in the IEEE 802.16 standard for real-time and non-real-time services. This strategy leads to more precise estimates of the amount of resources that should be allocated for each admission decision.

Although Chandra and Sahoo [6] consider the overhead of different types of service, their mechanism does not include the extended real-time polling service (ertPS) and it does not provide methods for the maximization of revenue and utility, which are important in the successful deployment of IEEE 802.16 networks on a commercial scale. Rong *et al.* [7] proposed a policy for revenue, utility, and fairness maximization. However, they assume that all connections using the same type of service have the same minimum traffic rate requirements, which does not comply with the IEEE 802.16 standards. The policies introduced in this paper do not make such assumption.

To the best of the authors' knowledge, this is the first paper to propose standard-compliant AC policies for revenue and utility maximization for IEEE 802.16 networks. This paper is a revised version of [8] containing enhanced mathematical formulation and numerical examples. The rest of this paper is organized as follows. Section 2 discusses related work. Sections 3 to 7 present the AC policies proposed in this paper. Section 8 describes the simulation environment used to evaluate these five policies. Section 9 presents numerical results and Section 10 concludes the paper.

2. RELATED WORK

Various proposals have been made for AC and pricing integration [9–14] as well as AC mechanisms for the maximization of utility/user satisfaction [15–18]. However, few address such issues in IEEE 802.16 networks.

Antonopoulos and Verikoukis [19] introduced an AC mechanism for the IEEE 802.16 standard, which provides higher priority for voice over IP (VoIP) calls than for other types of traffic in the network. When arrival rates of unsolicited grant service (UGS) connections are high, a UGS request is accepted if the total available bandwidth (BW_{T}) is sufficient to serve the incoming connection. In the case of ertPS, real-time polling service (rtPS), and non-real-time polling service (nrtPS) connections, restricted bandwidth is provided ($BW_{\mathrm{T}} - BW_{\mathrm{R}}$). Requests of best effort (BE) connections are always accepted. The portion of reserved bandwidth for UGS connections (BW_{R}) is dynamically adjusted according to the traffic intensity of the VoIP calls.

Yu *et al.* [20] introduced a statistical AC mechanism based on the complete sharing algorithm, which takes into account the overflow probability caused in the variability of the traffic and in the channel state of IEEE 802.16 networks. Connections transmitting variable bit rate traffic are accepted only if the requested bandwidth plus the variance of this requested bandwidth can be provided. In the formulation of the problem, the total bandwidth required by all accepted connections is assumed to follow a Gaussian distribution, which, according to the authors, involves a Gaussian process approximation that is limited.

Rong *et al.* [7] developed a utility-constrained and fairness-constrained optimal revenue AC policy as well as its corresponding approximation algorithm. Utility maximization consists of maximizing the amount of bandwidth allocated to the users, whereas fairness maximization consists of maintaining the blocking probabilities of all traffic classes relatively uniform. Nevertheless, the solution proposed in this paper is not standard-compliant, because it assumes that all connections using the same type of service have the same minimum traffic rate requirement.

The AC policies introduced in the following sections are standard-compliant and allow revenue and utility maximization while being simple to implement. Standard compliance and implementation complexity are major constraints when designing AC mechanisms for WiMAX networks [21].

3. STANDARD-COMPLIANT ADMISSION CONTROL POLICY

In this policy, the AC decision considers only the minimum rate requirement defined in the IEEE 802.16 standard. It accepts a new connection when the following condition is satisfied:

$$(C_{\mathrm{reserved}} + TR_{ij} \leq C) \qquad (1)$$

where TR_{ij} is the traffic rate that should be guaranteed for the new connection j of service type i; C_{reserved} is the capacity reserved for the existing connections, which is equal to $\sum_{i=1}^{4}\sum_{j=1}^{n_i} TR_{ij}$; and C is the capacity available for the uplink scheduler, that is, the amount of uplink bandwidth that the scheduler can allocate for data and bandwidth request transmissions.

A connection requesting admission informs, among other parameters, its *minimum reserved traffic rate* requirement ($minTR_{ij}$). However, this rate is not sufficient for the connection when the overhead generated by the type of service flow being solicited is considered. Therefore, instead of considering only the *minimum reserved traffic rate* requirement, the TR_{ij} value also includes an estimation of the bandwidth overhead generated by service flow i.

According to the IEEE 802.16 standard, UGS and ertPS connections receive grants at intervals defined by the *unsolicited grant interval* (ugi) parameter. The grant size for UGS connections is fixed, based on the *maximum sustained traffic rate* of the service flow (which is equal to the *minimum reserved traffic rate* for this service). Although the size of the grants allocated to ertPS connections varies according to the amount of requested bandwidth, the minimum size that should be guaranteed is based on the *minimum reserved traffic rate*. In this way, for both types of service flow, the base station should guarantee grants with size equal to $grantSize_{ij}$, where

$$grantSize_{ij} = minTR_{ij} * ugi_{ij} \qquad (2)$$

Equation (2) gives the grant size in bytes. However, grants are awarded in slots; thus, the number of slots assigned to a connection j approximates the smallest integer larger than or equal to $grantSize_{ij}/slotSize$. This normalization may result in bandwidth waste when $grantSize_{ij}$ is not a multiple of the number of bytes that can be transmitted in a single time slot ($slotSize$). Consequently, the rate that will be used by an UGS (ertPS) connection can be larger than the *maximum sustained traffic rate* (*minimum reserved traffic rate*) requirement. To account for this overhead, the value of TR_{ij} for UGS and ertPS connections is computed as follows:

$$TR_{ij} = \frac{\left\lceil \frac{grantSize_{ij}}{slotSize} \right\rceil * slotSize}{ugi_{ij} + tj_{ij}} \quad (3)$$

In Equation (3), the value of $\left\lceil \frac{grantSize_{ij}}{slotSize} \right\rceil$ gives the grant size in slots. Grants should be provided in intervals lasting ugi_{ij} with the maximum jitter tj_{ij} (defined by the parameter *tolerated jitter*). Multiplying the grant size in slots by the size of a single slot in bytes and dividing this value by the sum $ugs_{ij} + tj_{ij}$ gives the minimum rate that should be guaranteed for either UGS or ertPS connections.

In addition to the *minimum reserved traffic rate*, rtPS and nrtPS connections also need periodic grants to request bandwidth (unicast polling) [1]. The rate that should be guaranteed for these connections is equal to the sum of the $minTR_{ij}$ requirement and the rate used by unicast polling as follows:

$$TR_{ij} = minTR_{ij} + \frac{upSlots * slotSize}{upi_{ij}} \quad (4)$$

where $upSlots$ is the number of slots used by unicast polling and upi_{ij} is the value provided by the *unsolicited polling interval* parameter. nrtPS connections do not provide this upi_{ij} parameter; however, according to the standard, this interval can be chosen by the base station.

Best effort connections are always accepted, because no QoS requirements are involved.

4. OPTIMAL ADMISSION CONTROL POLICY FOR REVENUE MAXIMIZATION

The AC policy presented in this section extends the one presented in Section 3 so that provider revenue can be maximized.

As for the previous policy, before admitting a new connection, the AC verifies whether the minimum rate requirement of the new connection can be guaranteed without affecting the QoS provided for the connections already admitted into the network. The overhead incurred by each type of service is also considered.

To maximize the revenue, the admission controller needs to collect all admissions requested during a certain time interval in order to decide which connections should be accepted. An optimal policy is expected to choose the set of connections, which provides the greatest revenue subject to network capacity. This problem can be reduced to the Stochastic Knapsack problem [22] by mapping the network capacity for the knapsack as well as the minimum rate required by each connection for the size of each object that should be placed in the knapsack. The optimization problem is modelled as an ILP problem in the following way:

$$max \sum_{i=1}^{4} \sum_{j=m_i+1}^{n_i} x_{ij} r_i minTR_{ij}$$

Subject to:

$$x_{ij} = 1 \ \forall i \in \{1,2,3,4\} \ \forall j \in \{1,\ldots,m_i\} \quad (R1)$$

$$x_{ij} \in \{0,1\} \forall i \in \{1,2,3,4\} \ \forall j \in \{m_i+1,\ldots,n_i\} \quad (R2)$$

$$\sum_{i=1}^{4} \sum_{j=1}^{n_i} x_{ij}(minTR_{ij} + \epsilon_i) \le C \quad (R3)$$

where

- n_i is the total number of service-i connections (existing connections plus new connections soliciting admission);
- m_i is the number of existing service-i connections;
- $minTR_{ij}$ is the minimum traffic rate requirement for the connection j, where i is the type of service and j is the connection ID;
- ϵ_i is the overhead incurred by service i, such that $minTR_{ij} + \epsilon_i = TR_{ij}$, where TR_{ij} is the minimum rate that should be allocated to connection j given by Equation (3) for UGS and ertPS connections and by Equation (4) for rtPS and nrtPS connections;
- r_i is the revenue rate of a service-i connection;
- x_{ij} is the variable used to indicate whether a new connection should be admitted;
- C is the capacity available for the uplink scheduler.

In the proposed formulation, i varies from 1 to 4 to designate UGS, ertPS, rtPS, and nrtPS connections. The revenue obtained from the admission of a new connection is given by the function $r_i * minTR_{ij}$. According to this function, the greater the minimum rate required by the connection, the greater is the price the user has to pay for the service and, therefore, the greater is the revenue obtained from the admission of that connection. In this way, providers can guarantee price fairness among users of the same type of service, because the asked price is proportional to the minimum rate requirement. Restriction R1 stipulates that x_{ij} is equal to 1 if j is an existing connection. Restriction R2 determines that the value of the variable x_{ij} should be either 0 or 1 for new connections. A connection j is rejected when $x_{ij} = 0$ and accepted when $x_{ij} = 1$. Restriction R3 guarantees that the sum of the resources reserved for all connections does not surpass the capacity available for the uplink scheduler.

Since this problem can be reduced to the Stochastic Knapsack problem, which is known to be non-deterministic polynomial-time hard, it is unlikely that it can be solved in polynomial time. The next section introduces a computationally feasible solution to be implemented in WiMAX networks.

5. HEURISTIC FOR REVENUE MAXIMIZATION

The AC policy presented in this section uses a greedy approach to maximize the revenue. As for the policy introduced in Section 4, in the greedy approach, admission decisions are made in pre-defined intervals. Connections are admitted in decreasing order on the basis of revenue provided by each connection.

The Algorithm *MaxRevenueAC* presents the steps executed at every admission interval in order to decide which connections will be accepted. Admission requests sent by new connections are stored in a buffer. In the first step, the algorithm computes a revenue value for each new connection using the revenue function $r_i * minTR_{ij}$. Next, pending requests are sorted in decreasing order of potential revenue. In the following steps, the algorithm uses Equation (1) to decide which connections will be accepted, starting with the first element in the sorted buffer.

Next, the complexity of the proposed solution is investigated. Let l be the number of new connections requesting admission. There are $O(l)$ iterations in the algorithm; moreover, the operation in line 3 requires an average of $O(l \log l)$ sorting time when sorting methods such as quick-sort are used (with worst case proportional to l^2). Therefore, the complexity of the *MaxRevenueAC* algorithm is $O(l \log l)$.

ALGORITHM MAXREVENUEAC

1. **for** each connection j stored in the *buffer* **do**
2. $\quad revenue_{ij} \leftarrow r_i * minTR_{ij}$
3. sort the *buffer* following the decreasing order of the values of $revenue_{ij}$
4. **for** each connection j stored in the *buffer* **do**
5. $\quad$ **if** type of service of j is UGS or ertPS **then**
6. $\qquad$ compute TR_{ij} using Equation (3)
7. $\quad$ **else**
8. $\qquad$ compute TR_{ij} using Equation (4)
9. $\quad$ **if** $(C_{\text{reserved}} + TR_{ij} \leq C)$ **then**
10. $\qquad$ accept connection j
11. $\qquad$ $C_{\text{reserved}} \leftarrow C_{\text{reserved}} + TR_{ij}$
12. $\quad$ **else**
13. $\qquad$ reject connection j

6. OPTIMAL ADMISSION CONTROL POLICY FOR REVENUE AND UTILITY MAXIMIZATION

The AC policy proposed in this section attempts to maximize network revenue and utility. For service providers, revenue is a primary concern, whereas from the perspective of the users, the best AC policy is the one that can achieve maximum utility or, equivalently, maximum transmission rate.

The optimal revenue and utility strategy is introduced in the context of IEEE 802.16 networks as follows. Let r_i denote the revenue rate of a service-i connection. The revenue obtained with the admission of a new connection j is equivalent to the minimum rate requested by j and is given by the function $r_i * minTR_{ij}$. In the same way, let u_i denote the utility rate of a service-i connection. The utility gain function is then defined as $u_i * extraTR_{ij}$, where $extraTR_{ij}$ is the extra rate, beyond the requested minimum, to be allocated to connection j. The value of $extraTR_{ij}$ cannot surpass the difference $maxTR_{ij} - minTR_{ij}$, to prevent violation of the maximum rate requirement ($maxTR_{ij}$).

In order to maximize the revenue, the admission controller should collect requests during a given time interval to decide which ones should be accepted. For utility maximization, for each admission interval, the admission controller should decide which connections, either new or existing ones, should be granted additional resources. This optimization problem can be formulated as an MILP problem as follows.

$$max \sum_{i=1}^{4} \sum_{j=m_i+1}^{n_i} x_{ij} r_i minTR_{ij} + \sum_{i=1}^{4} \sum_{j=1}^{n_i} u_i extraTR_{ij}$$

Subject to :

$$x_{ij} = 1 \ \forall i \in \{1,2,3,4\} \ \forall j \in \{1,\ldots,m_i\} \quad \text{(R1)}$$

$$x_{ij} \in \{0,1\} \forall i \in \{1,2,3,4\} \forall j \in \{m_i+1,\ldots,n_i\} \quad \text{(R2)}$$

$$0 \leq extraTR_{ij} \leq maxTR_{ij} - minTR_{ij}$$
$$\forall i \in \{1,2,3,4\} \ e \ \forall j \in \{1,\ldots,m_i\} \quad \text{(R3)}$$

$$0 \leq extraTR_{ij} \leq (maxTR_{ij} - minTR_{ij}) x_{ij}$$
$$\forall i \in \{1,2,3,4\} \ e \ \forall j \in \{m_i+1,\ldots,n_i\} \quad \text{(R4)}$$

$$\sum_{i=1}^{4} \sum_{j=1}^{n_i} x_{ij} (minTR_{ij} + \epsilon_i + extraTR_{ij}) \leq C \quad \text{(R5)}$$

where

- n_i is the total number of service-i connections (existing connections plus new connections soliciting admission);
- m_i is the number of existing service-i connections;
- $minTR_{ij}$ is the minimum traffic rate requirement for the connection j, where i is the type of service and j is the connection ID;
- $maxTR_{ij}$ is the maximum traffic rate requirement for the connection j;
- ϵ_i is the overhead incurred by service i, such that $minTR_{ij} + \epsilon_i = TR_{ij}$, where TR_{ij} is the minimum rate that should be allocated to connection j given by Equation (3) for UGS and ertPS connections and by Equation (4) for rtPS and nrtPS connections;

- $extraTR_{ij}$ is the additional rate, beyond the minimum rate requirement, which will be allocated to the connection j;
- r_i is the revenue rate of a service-i connection;
- u_i is the utility rate of a service-i connection;
- x_{ij} is the variable used to indicate whether or not a new connection should be admitted;
- C is the capacity available for the uplink scheduler.

In the proposed formulation, i varies from 1 to 4 to designate UGS, ertPS, rtPS, and nrtPS connections. Restriction R1 stipulates that x_{ij} is equal to 1 if j is an existing connection. Restriction R2 determines that the value of the variable x_{ij} should be either 0 or 1 for new connections, with connection j being rejected when $x_{ij} = 0$ and accepted when $x_{ij} = 1$. Restriction R3 guarantees that the extra rate ($extraTR_{ij}$) allocated for an existing connection does not violate its maximum rate requirement. The amount of extra resources allocated to existing connections is updated at each admission interval in order to adapt it to the network dynamics. Restriction R4 guarantees that the extra rate allocated to a new connection is zero when the connection is rejected and that it does not violate the maximum rate requirement when the connection is admitted. It is important to note that this policy never allocates extra resources to UGS connections, because the minimum and maximum rate requirements are the same for this type of service. Restriction R5 guarantees that the sum of the resources reserved to all connections does not surpass the capacity available for the uplink scheduler.

The problem presented in this section is at least as hard as the the problem introduced in Section 4, because it requires satisfaction of additional constraints. Therefore, it is also unlikely that this AC policy can be optimally solved in polynomial time. The next section presents an algorithm to solve the AC problem for revenue and utility maximization in polynomial time.

7. HEURISTIC FOR REVENUE AND UTILITY MAXIMIZATION

The AC policy introduced in this section uses a heuristic to maximize both revenue and utility. As does the policy proposed in the previous section, this algorithm decides which connections should be admitted, as well as the amount of extra resources that should be reserved for both new and existing connections.

The algorithm is based on both cost and benefit of a decision. The benefit gained from the acceptance of a new connection j is equal to the sum of the revenue and the utility gain provided by it ($r_i * minTR_{ij} + u_i * extraTR_{ij}$). The cost of admitting a connection is equal to the amount of resources that should be reserved ($TR_{ij} + extraTR_{ij}$). Connections with demands that can be met are admitted in decreasing order in relation to the associated rewards (given by the ratio between the benefit and the cost).

For existing connections, the algorithm decides on the amount of extra resources that should be allocated to maximize utility. The benefits gained by allocating extra resources for an existing connection j is equal to $u_i * extraTR_{ij}$, whereas the cost is $extraTR_{ij}$. If the reward obtained by allocating extra resources for an existing connection is greater than that obtained by the admission of a new connection, the new connection is rejected and the existing connection receives extra resources.

The Algorithm *MaxRevenueUtilityAC* presents the steps taken at the beginning of each admission interval. First, the algorithm calculates the reward provided by new connections with values of TR_{ij} less than the available capacity and by existing connections not using the UGS service. New connections requesting more bandwidth than the available capacity are rejected, while existing UGS connections are not considered because they will not receive

ALGORITHM MAXREVENUEUTILITYAC

1. insert new and existing connections in the $reward\_array$
2. **for** each connection j in $reward\_array$ **do**
3. **if** j is a new connection **then**
4. **if** $TR_{ij} > C_{\text{available}}$ **then**
5. reject j and delete it from $reward\_array$
6. **else**
7. $extraTR_{ij} \leftarrow min(C_{\text{available}} - TR_{ij}, maxTR_{ij} - minTR_{ij})$
8. $benefit \leftarrow r_i * minTR_{ij} + u_i * extraTR_{ij}$
9. $cost \leftarrow TR_{ij} + extraTR_{ij}$
10. $reward_{ij} \leftarrow benefit \div cost$
11. **else**
12. **if** j is a UGS connection **then**
13. $extraTR_{ij} \leftarrow 0$
14. delete j from $reward\_array$
15. **else**
16. $extraTR_{ij} \leftarrow min(C_{\text{available}}, maxTR_{ij} - minTR_{ij})$
17. $benefit \leftarrow u_i * extraTR_{ij}$
18. $cost \leftarrow extraTR_{ij}$
19. $reward_{ij} \leftarrow benefit \div custo$
20. sort the the $reward\_array$ following the decreasing order of the values of $reward_{ij}$
21. **while** $C_{\text{available}} > 0$ **and** $reward\_array \mathrel{!}= \emptyset$ **do**
22. $j \leftarrow$ ID of the connection with the highest value for $reward_{ij}$
23. **if** $TR_{ij} > C_{\text{available}}$ **then**
24. reject j and delete it from $reward\_array$
25. **else**
26. **if** j is a new connection **then**
27. **if** $TR_{ij} + extraTR_{ij} <= C_{\text{available}}$ **then**
28. accept j with extra rate equal to $extraTR_{ij}$
29. $C_{\text{available}} \leftarrow C_{\text{available}} - (TR_{ij} + extraTR_{ij})$
30. delete j from $reward\_array$
31. **else** go to step 2
32. **else**
33. **if** $extraTR_{ij} <= C_{\text{available}}$ **then**
34. reserve to j extra rate equal to $extraTR_{ij}$
35. $C_{\text{available}} \leftarrow C_{\text{available}} - extraTR_{ij}$
36. delete j from $reward\_array$
37. **else** go to step 2

additional resources. To calculate the values of benefit and cost, it is necessary to define the amount of extra resources to be allocated for each connection ($extraTR_{ij}$). The algorithm sets $extraTR_{ij}$ to the maximum value possible. For new connections, this value is equal to the minimum between the available capacity after the connection is accepted ($C_{\text{available}} - TR_{ij}$) and the difference between the minimum and maximum rate requirements ($maxTR_{ij} - minTR_{ij}$). For existing connections, the maximum value for $extraTR_{ij}$ is equal to the minimum between the available capacity and the difference between the minimum and maximum rate requirements.

In the second step, the algorithm selects the connections providing the highest reward values. Whenever a connection is selected, the available capacity is updated by subtracting $TR_{ij} + extraTR_{ij}$ from $C_{\text{available}}$, for a new connection, or $extraTR_{ij}$, for an existing connection. When the available capacity is less than the amount that should be reserved for a selected connection, the algorithm returns to the first step to recalculate the $extraTR_{ij}$ values as well as the reward values for the remaining connections. The algorithm finishes when no available capacity remains or when all admission and allocation decisions have been taken.

Next, the complexity of the proposed solution is analyzed. Let n be the number of existing connections plus new connections soliciting admission. First, there are $O(n)$ iterations (lines 2–19) to calculate the reward value of each connection. Next, the sorting operation requires an average of $O(n \log n)$ sorting time. Finally, there are $O(n)$ iterations to decide which connections should be admitted as well as the allocation of extra resources, which, in the worst case scenario, may require the algorithm to go back to step 2 and recalculate the reward values. Therefore, the complexity of the *MaxRevenueUtilityAC* algorithm is $O(n^2 \log n)$.

8. SIMULATION EXPERIMENTS

To conduct this study, an ns-2 module for IEEE 802.16 networks, developed by the authors and publicly available, was used [23].

The simulated network consisted of a base station, with the subscriber stations uniformly distributed around it. The frame duration was 1 ms, and the capacity of the channel was 40 Mbps, assuming a 1:1 downlink-to-uplink time division duplex (TDD) split. Five different types of traffic were considered: voice, voice with silence suppression, video, File Transfer Protocol (FTP), and Web, which are associated with UGS, ertPS, rtPS, nrtPS, and BE services, respectively.

The voice model used was an exponential 'on/off' model with mean duration of the 'on' and of the 'off' periods equal to 1.2 and 1.8 s, respectively. During the 'on' periods, 66-byte packets are generated every 20 ms [24]. The voice with silence suppression model used the Enhanced Variable Rate Codec [25], with packets generated every 20 ms employing Rate 1 (171 bits/packet), Rate 1/2 (80 bits/packet), Rate 1/4 (40 bits/packet), or Rate 1/8 (16 bits/packet). Video traffic was generated by 10 different real MPEG traces [26]. The WEB traffic was modeled by a hybrid Lognormal/Pareto distribution, with the body of the distribution corresponding to an area of 0.88 modeled by a Lognormal distribution with a mean of 7247 bytes and the tail modeled by a Pareto distribution with mean of 10 558 bytes [27]. FTP traffic was generated using an exponential distribution with a mean of 512 KBytes.

The unsolicited grant interval for the UGS and for the ertPS service was 20 ms. The unsolicited polling interval for the rtPS service was 20 ms and for the nrtPS service was 1 s.

Each rtPS connection had its own minimum reserved traffic rate and maximum sustained traffic rate requirements, which varied according to the mean transmission rate of the video. The nrtPS service had a minimum reserved traffic rate requirement of 200 Kbps and a maximum sustained traffic rate requirement of 300 Kbps. The BE service did not have any QoS requirement.

The revenue rates were set as follows: 4 for the UGS service, 3 for the rtPS service, 2 for the ertPS service, and 1 for the nrtPS service. The same values were used for the utility rates, except that for the UGS service, which had utility rate set to 0, because connections using this type of service do not receive extra resources.

The lifetime of connections were exponentially distributed, with an average of 600 s for rtPS connections and 300 s for UGS, ertPS, and nrtPS connections. The connection arrival rates were governed by an exponential distribution with the mean varying from 60 to 2 s for each type of service. In this way, under the lowest load, an average of one connection of each type of service arrived per minute, whereas under the highest load, an average of 30 connections of each type of service arrived per minute.

Each result was produced by running the simulation 10 times using different seeds. The mean values and the 95% confidence intervals are shown in the figures.

9. NUMERICAL RESULTS

In this section, the policies proposed in Sections 3 to 7 are, respectively, referred as Simple AC, Optimal Max Revenue AC, Heuristic Max Revenue AC, Optimal Max Revenue and Utility AC, and Heuristic Max Revenue and Utility AC.

The proposed AC policies were tested in conjunction with the uplink scheduler proposed by the authors in [28] to evaluate their ability to support the scheduling mechanism in the provisioning of QoS. The uplink scheduler is able to guarantee the minimum traffic rate and maximum latency requirements when the network is not overloaded. For the simulation experiments carried out in this paper, the scheduler was adapted to perform the allocation of spare resources according to the decisions of the Optimal Max

Revenue and Utility AC, and Heuristic Max Revenue and Utility AC policies.

Figure 1 presents the blocking probability values for the proposed policies. When the network is overloaded, the Heuristic Max Revenue AC policy gives the highest blocking probabilities. For those scenarios, the difference between the blocking probabilities obtained with the optimal policy and the heuristic for revenue maximization can reach 13%. The Heuristic Max Revenue AC policy selects the connections to be admitted based on the decreasing order of revenue values associated with each connection. rtPS connections provide the greatest revenue values (computed as $r_i * minTR_{ij}$), because this service is associated with one of the highest rate of revenue return and due to the highest minimum rate requirement values; thus, the Heuristic Max Revenue AC policy accepts the highest number of rtPS connections. On the other hand, rtPS connections consume large amounts of resources, which results in high blocking probabilities for the UGS, ertPS, and nrtPS services, which explains the results in Figure 1.

Both the optimal policy and the heuristic for revenue and utility maximization provide identical values of blocking probability with both achieving the highest values when the arrival rate is five connections of each type of service per minute. Whereas the other three policies admit a high percentage of the requesting connections due to the low load in the network, the Optimal Max Revenue and Utility AC, and the Heuristic Max Revenue and Utility AC policies, which maximize the utility, reject a greater number of ertPS and nrtPS connections than do the other policies, due to the provision of larger amounts of bandwidth for admitted connections.

By separately analyzing the blocking probabilities of each type of service, it is possible to identify significant differences among them. Some types of service have more connections accepted than others to allow the AC policy to meet its goal. Figure 2, for example, presents the blocking probability for UGS connections. This type of service has the lowest blocking probability for all the proposed AC policies. The policies that maximize revenue and utility

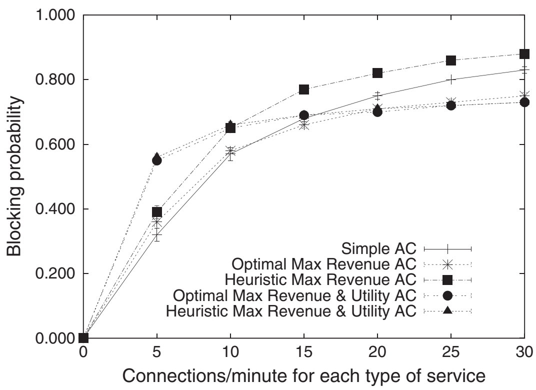

Figure 1. Blocking probability.

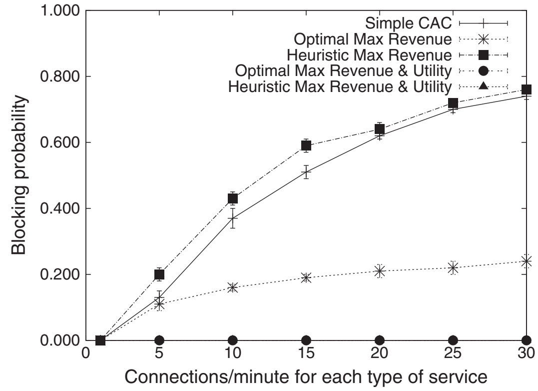

Figure 2. Lowest blocking probability for proposed AC policies.

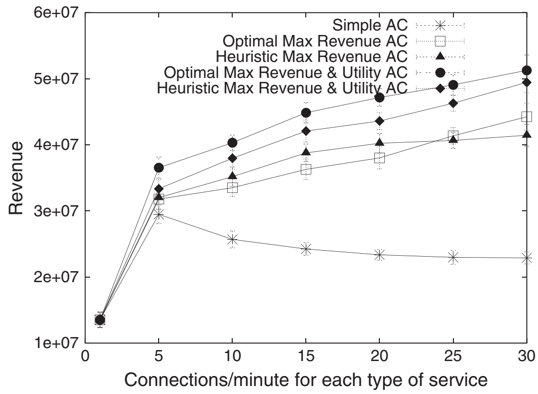

Figure 3. Revenue from Simple AC, Optimal AC, and Heuristic AC policies.

accept all UGS connections independent of the connection arrival rate.

These results show how policies with different goals may affect users of different types of service differently. When the network is overloaded, a service provider may wish to promote balanced blocking in order to avoid disparities between users of different types of service, even if, as a consequence, revenue is reduced. Such a policy can be accomplished by extending the proposed policies in order to minimize the difference between blocking probability values of the different types of service.

Figure 3 presents the accumulated revenue gain resulting from the proposed policies. As expected, the Simple AC policy generates the least revenue, because it is not concerned with such maximization. Moreover, because this policy considers only the minimum traffic rate requirement for the admission decisions, revenue values tend to stabilize as the network becomes overloaded. The greatest revenue with this policy is observed when the connections arrival rate is of five connections of each type of service per minute. In this scenario, as a result of the low load, the AC policy is able to admit the highest number of rtPS connections, which, as previously explained, provide the highest revenue values.

Both the optimal policy and the heuristic for revenue maximization produce similar revenue values. For scenarios with arrival rates of 10, 15, and 20 connections of each type of service per minute, the Heuristic Max Revenue AC policy yields slightly higher average revenue values than does the optimal approach. The solutions provided by the Optimal Max Revenue AC policy are locally optimal for each admission interval. However, local optimization methods do not guarantee a global maximum on the long run, so the decisions made by the heuristic at each admission interval led to higher average revenue values than the optimal approach in some scenarios.

The revenue obtained by the Heuristic Max Revenue and Utility AC policy is close to that produced by the optimal approach, regardless of the connection arrival rate; however, on average, the Optimal Max Revenue and Utility AC policy produces greater revenue than does the heuristic. Moreover, Figure 4 shows that the heuristic yields higher utility gain than does the optimal approach. These results can be explained by the fact that, although both policies produce similar blocking, the optimal approach admits an average of 1% to 2% more rtPS connections than does the heuristic algorithm. By admitting fewer rtPS connections, the Heuristic Max Revenue and Utility AC policy produces less revenue values than does the optimal approach, but it is able to allocate additional resources for the connections admitted leading to greater utility gains.

For both policies, the utility gain decreases as the arrival rate of connections increases. When the amount of available bandwidth is restricted, the admission of a new UGS connection is more profitable than allocating additional resources to the existing connections. According to the values configured for the revenue and the utility gains, the admission of a new UGS connection provides a gain of 4 units for each Kbps of minimum reserved traffic rate, whereas allocating additional resources for an rtPS connection provides a gain of 3 units for each Kbps of extra bandwidth. For the scenario where the connections arrival rate is of 1 connection per minute for each type of service, the number of admitted connections is low and, given the restriction on maximum rate, the AC policies are not able to allocate all the available bandwidth leading to a low gain in utility.

Latency values were collected for the UGS and the ertPS connections in order to check whether they receive the periodic grants specified by the standard. As can be seen in Figures 5 and 6, fluctuations on the latency values arising from load variation are in the order of a few milliseconds (note the small scale of the y-axis), which indicates that the scheduler was able to allocate the expected periodic grants for UGS and ertPS connections.

Results were also obtained to verify whether or not the average latency of the rtPS connections violated the maximum latency requirement. Figure 7 shows that the maximum latency requirement of 100 ms was guaranteed in all the simulated scenarios. Because the Simple AC policy does not maximize revenue, it blocks a large number of rtPS connections, thus leading to lower latency values than those obtained by the use of the other policies.

Furthermore, throughput values were verified to determine whether the minimum traffic rate requirements of rtPS and nrtPS connections were guaranteed.

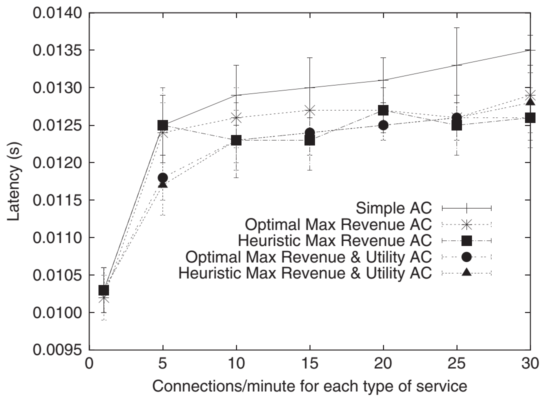

Figure 5. Average latency for unsolicited grant service connections.

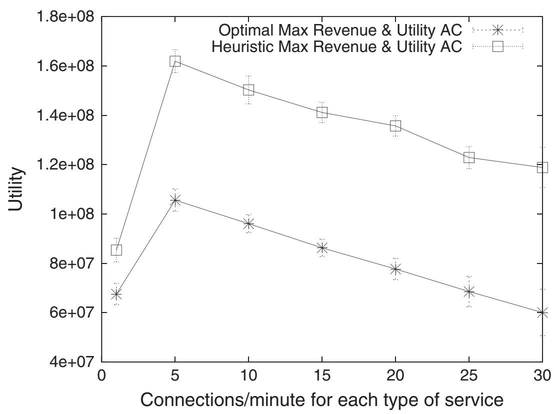

Figure 4. Utility of Optimal AC and Heuristic AC policies.

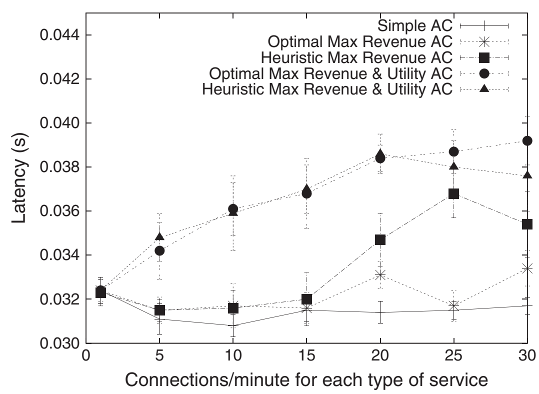

Figure 6. Average latency for extended real-time polling service connections.

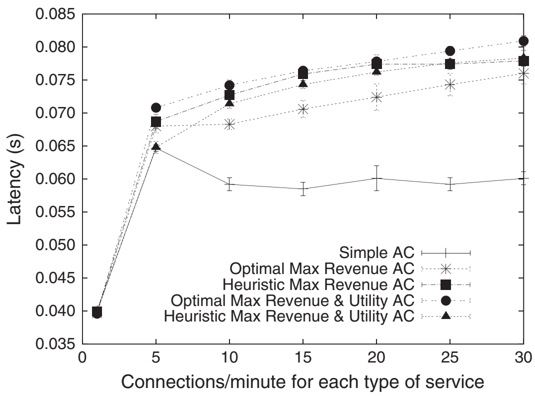

Figure 7. Average latency for real-time polling service connections.

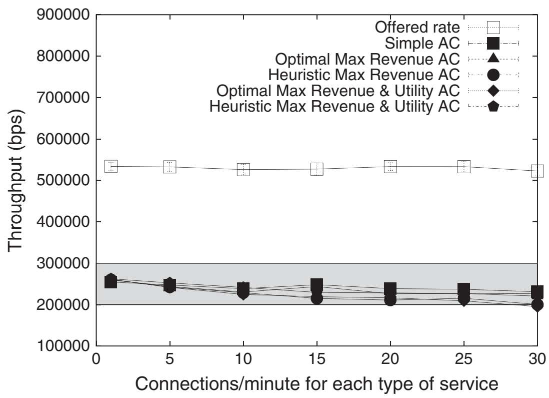

Figure 8. Average throughput of real-time polling service connections transmitting the Ice Age video.

Figure 8 shows the average offered rate as well as the throughput of the rtPS connections transmitting the Ice Age video. The minimum traffic rate requirement of 200 Kbps and the maximum traffic rate requirement of 300 Kbps were respected regardless of network load and AC policy being used.

The average throughput of nrtPS connections occurs in the interval defined by the requirement of the minimum traffic rate and the maximum traffic rates, as shown in Figure 9. Although the nrtPS flows were configured to generate an average rate equal to the maximum sustained traffic rate, the joint effect of control of maximum rate by the scheduler and TCP congestion control resulted in the offering load of less than 300 Kbps. The interaction between the scheduling mechanism and the TCP congestion control mechanism will be investigated in the future so that resource underutilization can be avoided. When the Simple AC policy was employed, all the nrtPS traffic generated by the upper layers was served at the MAC layer in all the simulated scenarios; consequently, the throughput of the nrtPS connections, measured at the MAC layer, was a little higher than the offered load, measured at the

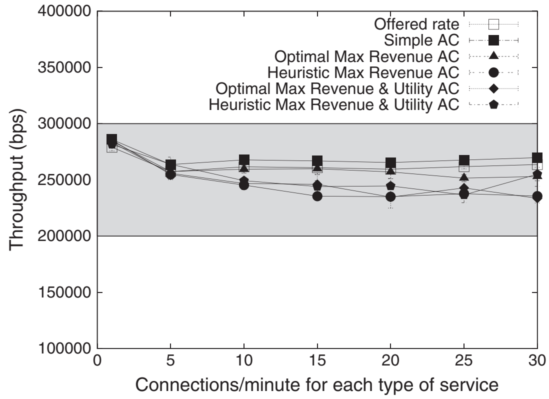

Figure 9. Average throughput of non-real-time polling service connections.

transport layer, because of additional load incurred by both MAC and network layer headers.

These simulation results show that the proposed policies are able to avoid saturation of the wireless channel so that the resources available to the uplink scheduler are sufficient to guarantee the QoS requirements of each admitted connection.

10. CONCLUSIONS

This paper introduced AC policies for IEEE 802.16 networks. The proposed policies restrict the number of simultaneous uplink connections in the system so that resources available to the uplink scheduler are sufficient to guarantee the QoS requirements of the connections. Moreover, these policies support the service provider and the users expectations by maximizing the network revenue as well as the network utility.

Five AC policies have been presented. First, a standard-compliant and easy-to-implement solution was introduced, it served as a basis for the development of the optimization policies. Two policies for the maximization of the revenue were proposed, one formulated as an ILP problem and the other one as a heuristic. Moreover, two other AC policies were introduced for revenue and utility maximization. The first one of these was formulated as an MILP problem, whereas the second used a heuristic.

Simulation results show that the proposed policies are able to fulfil the expectation of service provider and users as well as to support the scheduling mechanism for QoS provisioning as specified by the IEEE 802.16 standard. A service provider interested only in the standard functionalities of IEEE 802.16 networks could use the simplest proposed policy, which does not optimize revenue and utility, to provide the required QoS. In the case the service provider also wishes to maximize expectations, the heuristic proposed to maximize revenue and utility is an attractive option, because it has lower complexity than the optimal policy and it provides near-optimal solutions. In fact, in order to service providers sustain their competitiveness in the marketplace, revenue and utility are key points to be considered.

REFERENCES

1. IEEE Standard 802.16-2004, Part 16. *Air Interface for Fixed Broadband Wireless Access Systems* 2004.
2. Wang H, He B, Agrawal DP. Above packet layer level admission control and bandwidth allocation for IEEE 802.16 wireless MAN. *Simulation Modeling Practice and Theory* 2007; **15**(14): 266–382.
3. Chen J, Jiao W, Wang H. A service flow management strategy for IEEE 802.16 broadband wireless access systems in TDD mode, In *Proceedings of the IEEE ICC*, Seoul, Korea, 2005; 3422–3426.
4. Jakkakorpi J, Sayenko A. Measurement-based connection admission control methods for real-time services in IEEE 802.16e, In *Proceedings of the CTRQ'09*, Colmar, France, 2009; 37–41.
5. Wang L, Liu F, Ji Y, Ruangchaijatupon N. Admission control for non-preprovisioned service flow in wireless metropolitan area networks, In *Proceedings of the 4th European Conference on Universal Multiservice Networks*, Toulouse, France, 2007; 243–249.
6. Chandra S, Sahoo A. An efficient call admission control for IEEE 802.16 networks, In *Proceedings of the 15th IEEE Workshop on Local & Metropolitan Area Networks*, Princeton, USA, 2007; 188–193.
7. Rong B, Qian Y, Lu K, Chen H, Guizani M. Call admission control optimization in WiMAX networks. *IEEE Transactions on Vehicular Technology* 2008; **57**(4): 2509–2522.
8. Borin JF, da Fonseca NLS. Admission control policies for revenue and utility maximization in IEEE 802.16 networks, In *Proceedings of the IEEE Globecom*, Anaheim, USA, 2010; 1–5.
9. Kikilis AA, Ratsiatos SS, Rouskas AN. Game theoretical formulation of admission control and pricing in wireless networks: the case of cooperating providers, In *Proceedings of the IEEE Globecom*, San Francisco, USA, 2006; 1–5.
10. Lindemann C, Lohmann M, Thümmler A. Adaptive call admission control for QoS/revenue optimization in CDMA cellular networks. *Wireless Networks* 2004; **10**(4): 457–472.
11. Tong H, Brown T. Adaptive call admission control under quality of service constraints: a reinforcement learning solution. *IEEE Journal on Selected Areas in Communications* 2000; **18**(2): 209–211.
12. Hou J, Yang J, Papavassiliou S. Integration of pricing with call admission control to meet QoS requirements in cellular networks. *IEEE Transactions on Parallel and Distributed Systems* 2002; **13**(9): 898–910.
13. Ni J, Tatikonda S. Revenue optimization via call admission control and pricing for mobile cellular systems, In *Proceedings of the IEEE ICC*, Seoul, Korea, 2005; 3359–3364.
14. Yilmaz O, Chen I-R. Utilizing call admission control for pricing optimization of multiple service classes in wireless cellular networks. *Computer Communications* 2009; **32**(2): 317–323.
15. Piamrat K, Ksentini A, Viho C, Bonnin J-M. QoE-aware admission control for multimedia applications in IEEE 802.11 wireless networks, In *Proceedings of the IEEE 68th Vehicular Technology Conference*, Calgary, Canada, 2008; 1–5.
16. Lu N, Bigham J. On utility-fair bandwidth adaptation for multi-class traffic QoS provisioning in wireless networks. *Computer Networks* 2007; **51**(10): 2554–2564.
17. Zhang Y, Dai S, Zhou C, Li L, Li B. Bandwidth allocation scheme and call admission control for adaptive multimedia services in wireless cellular networks—part A: numerical results, In *Proceedings of the International Conference on Communications, Circuits and Systems*, Hong Kong, China, 2005; 346–352.
18. Wu Z, Yin Q. A heuristic for bandwidth allocation and management to maximize user satisfaction degree on multiple MPLS paths, In *Proceedings of the 3rd IEEE Consumer Communications and Networking Conference*, Las Vegas, USA, 2006; 35–39.
19. Antonopoulos A, Verikoukis C. Traffic-aware connection admission control scheme for broadband mobile systems. *IEEE Communications Letters* 2010; **14**(8): 719–721.
20. Yu K, Wang X, Sun S, Zhang L, Wu X. A statistical connection admission control mechanism for multiservice IEEE 802.16 network, In *Proceedings of the IEEE 69th Vehicular Technology Conference*, Barcelona, Spain, 2009; 1–5.
21. Msadaa IK, Camara D, Filali F. Scheduling and CAC in IEEE 802.16 fixed BWNs: a comprehensive survey and taxonomy. *IEEE Communications Surveys and Tutorials* 2010; **12**(4): 459–487.
22. Ross KW, Tsang DHK. The stochastic knapsack problem. *IEEE Transactions on Communications* 1989; **37**(7): 740–747.
23. Borin JF, da Fonseca NLS. Simulator for WiMAX networks. *Simulation Modeling Practice and Theory* 2008; **16**(7): 817–833.
24. Brady P. A model for generating on-off speech patterns in two-way conversations. *Bell System Technical Journal* 1969; **48**: 2445–2472.
25. 3GPP2 C.S0014-0, Enhanced Variable Rate Codec (EVRC).

26. Seeling P, Reisslein M, Kulapala B. Network performance evaluation using frame size and quality traces of single-layer and two-layer video: a tutorial. *IEEE Communications Surveys and Tutorials* 2004; **6**(2): 58–78.
27. Barford P, Bestavros A, Bradley A, Crovella M. Changes in web client access patterns: characteristics and caching implications. *Tech. Report 1998-023*, Boston Uni., 1998.
28. Borin JF, da Fonseca NLS. Scheduler for IEEE 802.16 networks. *IEEE Communications Letters* 2008; **12**(4): 274–276.

AUTHORS' BIOGRAPHIES

Juliana Freitag Borin is a professor at the Institute of Computing of the University of Campinas, Campinas, Brazil. She received her BSc in Computer Science (2001) from the University of Western Paraná, Brazil, and her MSc (2004) and PhD (2010) degrees in Computer Science from the University of Campinas. After her PhD, she worked one year as a PostDoc at the University of Campinas. She is an expert on Broadband Wireless Networks architecture, design, quality-of-service provisioning, and performance and has published several research papers in the area.

Nelson Luis Saldanha da Fonseca received his Electrical Engineer (1984) and MSc in Computer Science (1987) degrees from The Pontificial Catholic University at Rio de Janeiro, Brazil, and the MSc (1993) and PhD (1994) degrees in Computer Engineering from The University of Southern California, USA. Since 1995, he has been affiliated with the Institute of Computing of The State University of Campinas, Campinas–Brazil, where he is currently a Full Professor. He published over 300 refereed papers and supervised over 50 graduate students. He served as Editor-in-Chief of the IEEE Communications Surveys and Tutorials, and Editor-in-Chief of the IEEE Communications Society Electronic Newsletter, and Editor of the Global Communications Newsletter. He is a member of the editorial board of: Computer Networks, IEEE Communications Surveys and Tutorials, IEEE Communications Magazine, Peer-to-Peer Networking and Applications, International Journal of Communication Systems, and Journal of Internet Service and Applications. He served on the editorial board of the IEEE Transactions on Multimedia, Brazilian Journal of Computer Science, and on the board of the Brazilian Journal on Telecommunications. He is the recipient of Elsevier Computer Networks Editor of the Year 2001. Currently, he is ComSoc Vice-president Member Relations and served as Director for Latin America and as Director of On-line Services.