

Congestion Avoidance Control through Non-cooperative Games between Customers and Service Providers

Dimitris E. Charilas, Athanasios D. Panagopoulos, Panagiotis Vlacheas,
Ourania I. Markaki, and Philip Constantinou

National Technical University of Athens, Department of Electrical & Computer Engineering,
Heron Polytechneiu 9, Zographou, 15773, Athens, Greece
dcharilas@mobile.ntua.gr, thpanag@ece.ntua.gr,
panvlah@telecom.ntua.gr, omarkaki@epu.ntua.gr,
fkonst@mobile.ntua.gr

Abstract. Congestion avoidance control refers to controlling the load of the network by restricting the admission of new user's sessions and resolving the unwanted overload situations. Admission control and Load control constitute key mechanisms regarding Radio Resource Management. As the wireless world is moving towards heterogeneous wireless networks, these types of control are facing more challenges, since efficiency and fairness are required. Game theory provides an appropriate framework for formulating fair and efficient congestion avoidance control problems. In this paper we formulate a non-cooperative game between service providers and customers. On the one hand, the service providers wish to maximize their revenue, but on the other hand, the users wish to maximize the quality of service received, keeping at the same time the expenses as low as possible. Therefore a balance has to be established among these contradictory demands. Our effort also concentrates in the proper modeling of the user's level of satisfaction, so as to provide a logical decision-taking framework. The proposed scheme is then tested using the ns2 simulator. Results show that both parties can benefit from this mechanism.

Keywords: Congestion Avoidance, Admission Control, Load Control, Quality of Service, Game theory, Payoff, Non-cooperative games, Nash Equilibrium.

1 Introduction

Wireless networks have limited radio resources and should be managed very carefully so that they operate under normal conditions, thus assuring that the users will receive the requested Quality of Service (QoS) for their requested applications/services. To retain its customer base, the service provider must make sure that customers are satisfied with the level of QoS they receive, taking into account the premium they pay. The level of customer satisfaction received can be represented by *utility*-based functions, due to the fact that each customer spends his/her disposable income in the way that yields him/her the greatest amount of satisfaction. This leads to the maximization of the utility functions [1]-[5]. Because of the limited radio resources, there have been defined Radio Resource strategies in order not only to assure the QoS guarantees to

the users, but also to assure that the users won't violate these agreements from their side. When, for example, there are too many users admitted and they don't receive the agreed QoS, the network is overloaded. This state of the network constitutes the congestion situation. Congestion situations are very harmful for the network because they cause many problems, such as increased interference, loss of packets, low bandwidth availability and from the user's point of view it causes decreased QoS reception, which leads to the user's disappointment.

The **Congestion Avoidance Control** mechanism [2] consists of Admission control (AC) and Load Control (LC). **Admission Control** is one of the key Radio Resource Management (RRM) mechanisms that ensure the proper operation of a network, by admitting or rejecting new user requests based on criteria such as the load of the network. In general, the Admission Control mechanism ensures that the admittance, of a new flow into a resource-constrained network, does not violate the QoS commitments already made by the network to the admitted flows. **Load Control** is also one of the key RRM mechanisms that serve for the effective performance of a wireless network by keeping the load of the network at normal boundaries. It performs traffic balancing between nodes or cells of the same mode preventing congestion situations. Reactive load control is employed to encounter overload situations of the network, when the users' QoS is at high risk. In these cases, the load control performs several actions to decrease the amount of traffic in the congested cell. These unwanted congested states may be prevented by the load control mechanism that monitors continuously the system.

On the other hand, **Game theory** is a mathematical tool developed to understand competitive situations in which rational decision makers interact to achieve their objectives. Game theory techniques have recently been applied to various engineering design problems in which the action of one component impacts (and perhaps conflicts with) that of any other component. In [6] the authors review popular game theory techniques, with regard to, the wireless networks' resource management problem and propose a game theoretic framework for optimizing bandwidth allocation and admission control issues in wireless networks. Existing game theory-based approaches to wireless networks' resource management consider that the game is played either among users, competing for network resources, or among networks, which try to maximize their efficiency by serving the largest possible number of available service requests. A different approach is adopted in [7],[8], where the AC problem is formulated as a non cooperative, non zero-sum game between the service provider and the customers, so as to increase the provider's revenue and offer differentiated QoS to the users. The authors of [7] also provide the required framework for n-player games. Finally, in [4], [9], the network selection problem is modeled by defining a game between the access networks involved in 4G converged environments. Decisive factor for the admission of service requests in all afore-mentioned works is the maximization of the payoff. The utility of the players is taken as the combination of strategies chosen in the game [10]. In this paper we intend to expand the methodology presented in [8], in order to include the efficient load control and guaranteed congestion avoidance mechanisms. The proposed scheme is then tested through extended simulations.

The paper is organized as follows. Section 1 is the introduction to the problem that we are dealing with. Sections 2 and 3 model the AC and LC games respectively. Section 4 shows the results deriving from simulations conducted on ns2. Finally, Section 5 summarizes the work.

2 The Admission Control Game

In this section we analyze the competitive customer vs. provider scenario as a non-cooperative two-player game. In the proposed scheme we consider that each customer has a contract with a specific service provider, thus him being the default network choice (“home” provider); nevertheless, in case of insufficient resources, the customer is free to pursue higher QoS at another provider, given that there is some kind of federation agreement between the visited and the home provider as in the roaming scenario (possibly under a small monetary penalty). Suppose that there are N users and M service providers, which means that each user at any time can choose any provider, giving a total of M^N possible states. Also, let $n_i(t)$ be the number of users subscribed to provider i at time t , $1 \leq i \leq M$. Furthermore, we assume that the user is not allowed to be subscribed simultaneously to multiple providers, meaning $\sum_{i=1}^M n_i(t) = N$. Each user-provider combination is considered as a two-player game G_j , $1 \leq j \leq M$.

Admission control takes place each time a new session request is received and decides whether it should be allocated resources or be rejected due to lack of resources. The decision is based on measurements extracted from on-going sessions of the same service type. Therefore, each time a new request is made, an instance of the game is played, as depicted in Figure 1. We assume that the service provider has two choices: either admit (S_1) or reject (S_2) the request. The customer also possesses two strategies: either leave (C_1) or stay (C_2) with the service provider, leaving us with four possible strategy combinations. The payoffs of the two players are expressed by the matrices $A = [a_{ij}]_{2 \times 2}$ and $B = [b_{ij}]_{2 \times 2}$. Table 1 presents schematically the relationships between payoffs and player strategies.

Table 1. Relationships between payoffs and player strategies

	Customer Leaves (C_1)	Customer Stays (C_2)
Provider Admits (S_1)	a_{11}, b_{11}	a_{12}, b_{12}
Provider Rejects (S_2)	a_{21}, b_{21}	a_{22}, b_{22}

Assume that a customer requests a session admission. We define his payoff matrix $B = [b_{ij}]_{2 \times 2}$ as follows:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} w_1 R - w_2 L_c & R \\ w_1 R_o - w_2 L_c & R_o \end{pmatrix} \quad (1)$$

Let us explain these payoff values.

- The term b_{21} denotes the case in which the user decides to leave, while the provider chooses not to admit access (strategy S_2C_1). In this case the user has a certain revenue R_o and is obliged to pay a penalty L_c for early termination, given that he has a contract with the service provider. Both these terms are multiplied by weights (w_1 and w_2 respectively), which reflect the user’s preference to save money and satisfaction respectively. The values of these weights may be specified through the user profile and indicate whether the user is risk neutral, risk seeking or risk adverse.

- Similarly, b_{22} describes the state where the service provider denies admission but the user chooses to stay, therefore having only the R_o revenue. Note that R_o is not calculated similarly to R , since no service is in progress. Its value may be fixed.
- The term b_{12} is defined as the user's revenue, in case he chooses to stay and he is granted admission.
- Finally, b_{11} is simply defined as the revenue minus the penalty, both multiplied by the weights w_1 and w_2 respectively.

The user's revenue expresses in monetary value the quality of service offered to him, taking into account the cost, and is modeled as

$$R = QoS(\%) \cdot q - C_{cust} \quad (2)$$

where q is a constant factor mapping the $QoS(\%)$ value to monetary value (specified by the provider), C_{cust} is the cost of the service from the customer's point of view and $QoS(\%)$ is given by (3). This difference between the monetary value of the QoS offered to the user and the actual price charged is known in microeconomic terms as the Consumer Surplus.

$$QoS(\%) = \sum_{i=1}^4 w_i \cdot Parameter_i \% \quad (3)$$

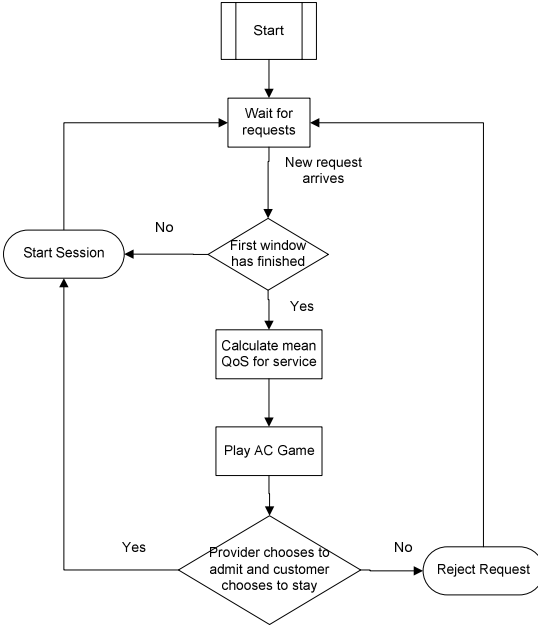
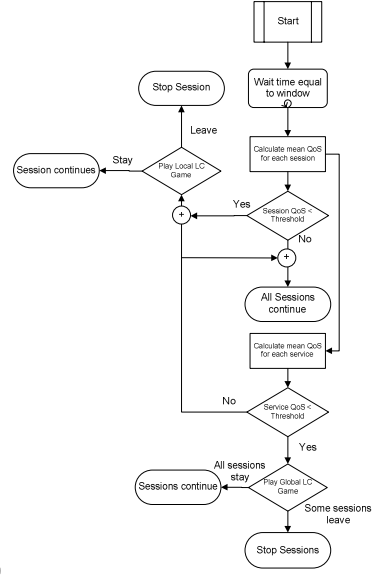
QoS(%) expresses a percentage of user satisfaction, taking into consideration the normalized mean values of QoS parameters such as Delay, Jitter, Throughput and Packet Loss. The weights w_i , $1 < i < 4$, vary with respect to the service and can be calculated based on the network's performance, as explained in [11][12]. The reason for this choice is because user satisfaction is subjective and therefore difficult to characterize mathematically. As a result, for the purposes of this work, we will use (3) to estimate the level of each customer's satisfaction or dissatisfaction, taking into account, four QoS parameters instead of only Call Blocking Probability, as in [7]. We consider this approach more concrete since it reflects the network's current status and is not dependent on constant variables, whose optimal values need to be specified.

Whenever the AC Game is performed, statistics are extracted from the data sent or received during a specified time interval. The duration of this interval is referred to as a "window". Therefore, the QoS assigned to a session or service will indicate only the most recent state information. Note that in case of requests during the first window, the AC Game is not performed since the window is not yet completed, meaning that these requests will always be granted admission. Now we define the provider's payoff matrix $A = [a_{ij}]_{2 \times 2}$ as:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} R_i + L_c + C - F - L & R_i + C - F \\ R_i + L_c - L & R_i \end{pmatrix} \quad (4)$$

Let us explain these payoff values as well.

- The term a_{22} represents the total revenue of the provider, deriving from all ongoing sessions. In other words, it shows what the provider is gaining in strategy S_2C_2 .
- Similarly, the term a_{21} shows the provider's gain if the customer chooses to leave, therefore we need to subtract the loss L from the total revenue and add the penalty L_c .


Fig. 1. Admission Control (AC) Game

Fig. 2. Load Control (LC) Game

• The term a_{12} denotes the provider's payoff in strategy S_1C_2 . This seems to yield the highest payoff for the provider. However, in a fully loaded system admitting a new customer may reduce the resources offered to on-going sessions (which means QoS degradation), thus causing other customers to leave and consequently leading to revenue loss. This degradation can be specified through real-time monitoring in simulations. Therefore, the total revenue equals the revenue gained from on-going sessions C plus the expected revenue from the new session minus the potential loss due to the dissatisfaction of other customers. The latter is expressed by the term F , in which its calculation will be discussed later.

• In a similar way, the term a_{11} shows the provider's revenue in case the customer chooses to **leave**. The term L corresponds to the revenue loss, while L_c corresponds to the penalty the customer has to pay for leaving.

Assuming the fact that the more the user is satisfied, the less likely he is going to leave and since $QoS(\%)$ expresses the user's satisfaction, we can estimate the probability that one customer leaves his current provider through the following equation

$$p_{leave} = 1 - (QoS(\%) / 100) \quad (5)$$

This allows us to express F in monetary value as the sum of all possible losses from all current customers, as

$$F = \sum_{i=1}^{n(t)} p_{leave_i} L_i \quad (6)$$

where p_{leave_i} is the probability that customer i chooses to leave and L_i is the corresponding revenue loss of the network provider.

As far as the solution of the game is concerned, two cases are distinguished. Assuming the case where the system is not full, the user request will be accepted and the probability that a customer leaves p_{leave_i} is near to 0. In this case, there is a Nash equilibrium at strategy pair S_1C_2 , that means the service provider accepts the request while the user remains with the provider. Assuming now the case where the system is loaded to a certain extent or even overloaded, the user request may be not accepted and the probability that a customer leaves p_{leave_i} is non zero. Even in this case, there is also a pure strategy Nash equilibrium at the pair S_1C_j , which depends on the relation between some terms in the payoffs. The new request is accepted if the revenue C generated from admitting the request is greater than the possible revenue loss F if the user leaves. Otherwise, the provider is better to reject the request. Based on these two proofs, a pure strategy Nash equilibrium is guaranteed and a solution to the admission control game is available at any point.

$$i = \begin{cases} 1 & \text{if } a_{11} \geq a_{21} \\ 2 & \text{if } a_{11} < a_{21} \end{cases} \Rightarrow i = \begin{cases} 1 & \text{if } C \geq F \\ 2 & \text{otherwise} \end{cases} \quad (7a) \quad j = \begin{cases} 1 & \text{if } \{i=1 \text{ and } b_{11} \geq b_{12}\} \text{ or } \{i=2 \text{ and } b_{21} \geq b_{22}\} \\ 2 & \text{if } \{i=1 \text{ and } b_{11} < b_{12}\} \text{ or } \{i=2 \text{ and } b_{21} < b_{22}\} \end{cases} \quad (7b)$$

3 The Load Control Game

The **Load Control (LC) Game** is played similarly to the AC Game. The main difference is that the LC Game is played periodically while the sessions are running. Through this process we intend to terminate sessions that greedily consume the system's resources, causing this way degradation to the QoS offered to the rest of the customers and thus reducing the provider's total revenue. Moreover, unsatisfied customers are granted the opportunity to seek more efficient networks, based on their preferences. Risk adverse customers for example will tolerate low levels of QoS and prefer to stay with the same provider in order to minimize the total cost.

Two types of LC Games are distinguished: the Local LC Game and the Global LC Game. Figure 2 shows the Load Control process. As mentioned before, the LC process is repeated periodically, independently from the AC process. We refer to the time interval between two successive rounds as a "**window**", since the LC Game will consider only the data received or sent during the last window. Therefore, the QoS assigned to a session or service will indicate only the most recent condition, enabling the LC scheme to react quickly in case of QoS degradation. In the beginning of every load control round, the QoS level for each session is extracted; then the average QoS level is specified for each service type. The QoS threshold is the lowest level of QoS that can be tolerated.

If the QoS of at least one service type is found below the acceptance threshold, then the **Global LC Game** is triggered, during which LC Games are played between the provider and all running sessions. This game may result in either disappointed customers leaving the provider or the provider terminating unprofitable customers. If either one decides that a connection should be terminated, then the session ends and the customer is prompted to another service provider. Penalty is submitted only if the customer chooses to leave willingly. On the other hand, if the Global LC Game is not

triggered and at least one session presents a QoS below the acceptance threshold, then the **Local LC Game** is triggered. In this case, an LC Game is played between the provider and each session that triggered the game, leaving the other sessions unaffected. This type of game may also result in some sessions being terminated.

4 Simulation Results

The proposed scheme was tested on the ns2 platform, using both CBR and TCP traffic. More specifically, we consider the voice service as a CBR/UDP service, defined by the G.711 codec, setting packets of 120 bytes and interval equal to 15 ms. Secondly, we consider the FTP/TCP service with packets of 512 bytes. Whenever a game is called, the trace files produced by ns2 are processed with the help of an awk script and the mean values for QoS parameters are extracted.

For the purposes of this study we have adopted the parameter values shown in Table 2. Parameters C_{cust} , L_c , c , L and R_o express monetary value (for example euros or American dollars), referring to the total duration of a session. These values may be derived from the provider's statistics. On the other hand, q refers only to the duration of a single time window and its value should be specified properly. In other words, q indicates the monetary value of a session with 100% QoS for a time equal to the duration of a window. Note that based on the selected values for C_{cust} , L_c and q , the QoS(%) threshold is specified as 40% for the voice service and 70% for the FTP service. The QoS percentage is estimated through the normalization of current parameter values in the interval specified by a minimum and maximum value, indicating the system's worst and optimal performance respectively, as shown in Table 3. This should be taken into account during the explanation of results, since different normalization intervals differentiate the same QoS percentage in each service.

Furthermore, in the frame of ensuring the best possible QoS for CBR connections, a new FTP request is allowed to play the AC game only if the mean QoS for all ongoing voice sessions is over 60%. The service mix was set to 2:1, while the simulation duration was set to 120 seconds. This time has been proven to be enough for the system to be congested at least once. Also note that, for simplicity reasons, the values of weights w_1 and w_2 were considered the same for all requests. In reality, those values should be different for each user in order to reflect his actual preferences.

Table 2. Simulation parameters

Q	L_c	C_{cust}	C	L	R_0	W_1	W_2	window
0.03	0.3	1.32 for Voice 2.4 for FTP	3	1	0.1	0.6	0.4	10 sec

Table 3. Normalization values

	Voice			FTP		
	Min	Max	w	Min	Max	w
Delay(ms)	10	60	0.4	10	50	0.2
Jitter(ms)	0	10	0.4	0	20	0.1
Thr/put(kbps)	10	65	0.1	10	100	0.7
P. Loss (%)	0	1	0.1	0	1	0.0

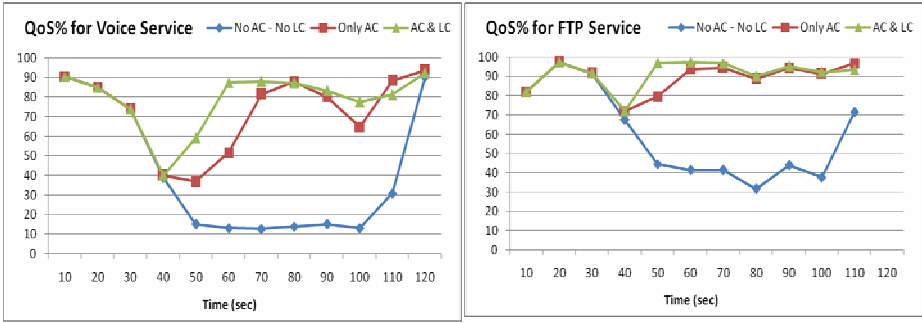


Fig. 3. Impact of AC and LC Games to QoS(%) a) Voice b) FTP

In order to monitor the improvement to offered QoS, we have implemented three tests. Firstly, we do apply neither AC nor LC mechanisms, meaning that all requests are accepted and no sessions are interrupted. Secondly we apply only the AC scheme and finally we apply both AC and LC schemes. In all cases, QoS is recorded for both service types at the end of each time window. Simulation results are depicted in Figure 3, where it can be seen that the third test provided the most satisfactory results. As expected, the absence of all kinds of congestion avoidance mechanisms resulted in major congestion, leaving the system unable to serve all sessions and thus resulting in unacceptable QoS levels. It can be easily observed that approximately after 40 seconds the system is congested. During the following time window, incoming requests are rejected, while load control terminates certain sessions as well. The system then recovers until a minor congestion takes place, approximately at $t=100$ seconds.

The main goal of this paper is to provide a scheme that maximizes not only the QoS offered to customers, but also the provider's gain. So far we have proven that the first part of this goal is indeed achieved. Therefore it is essential to examine whether the proposed scheme is actually in the provider's best interest as well. It is assumed that the provider's billing scheme takes into account the QoS percentage offered to customers, meaning that the customer pays an amount proportional to the level of QoS he receives. In this way, the provider has interest in optimizing the balance between the number of handled requests and QoS offered, instead of only maximizing the number of accepted requests. In conclusion, the total revenue deriving from all on-going sessions during a time window is shown by (8)

$$R_{total} = \sum_{i=0}^{i=service_num} \sum_{k_i=0}^{k_i} (QoS_{k_i} (\%) / 100) \cdot w_{cost} \quad (8)$$

where *service_num* indicates the number of service types, *k* the number of on-going sessions for the specific service type and w_{cost} the per window cost of a 100% QoS served session. In other words, w_{cost} constitutes the maximum amount of money the customer may pay during a single time window. As the QoS received decreases, so does the amount of money the customer is obliged to pay. Figure 4 shows the total revenue that derives from the three tests described in the previous section. Similarly

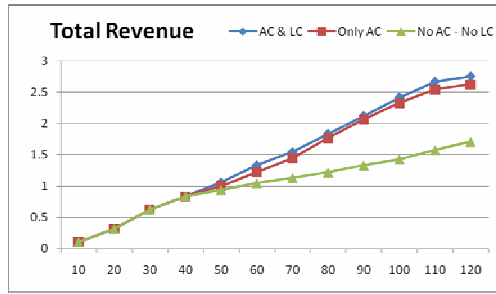


Fig. 4. Total Revenue for the provider

to QoS, the combination of AC and LC indeed offers the best revenue to the provider, thus this mechanism protects the interests of both parties.

The choice of window duration is considered of high importance, since an extremely high value may not allow the system to react quickly in case of congestion, while an extremely low value may add significant computational load and thus reduce the system's performance. When the window is set to a high duration, the system's reaction to the congestion is delayed, which lowers the quality of voice sessions. This occurs due to the fact that the system is unable to instantly detect congestions; therefore more requests are admitted in comparison with a smaller window. On the other hand, when the time window is set to a small duration, the system becomes extremely "strict" and results in rejecting voice requests which after a few seconds could be appropriately served. The window size should be chosen according to the expected average number of on-going sessions, since an increasing number of sessions also increases complexity and the algorithm becomes more time-consuming. Ideally, the window size should be given an initial value and dynamically adapt based on the system's status.

5 Conclusions

In this paper a methodology for integrating game theory in congestion avoidance mechanisms has been presented. The problem is seen as follows: the goal of service providers is to maximize their revenue without congesting their networks, though the goal of the customers is to get the maximum QoS with the minimum paying. So a congestion avoidance mechanism should take into account these factors in its decisions in order to maximize the satisfaction of both groups and simultaneously not congest the wireless networks. In this paper, the previously described problem has been modelled as a non-cooperative game between service providers and customers. Emphasis has also been given, on how we perceive and estimate the user's satisfaction based on network measurements. This approach provides a more subjective framework for the estimation of QoS as it is perceived by the end user. Finally, the simulations of the proposed scheme are presented and some useful conclusions are drawn.

References

- [1] Das, S.K., Lin, H., Chatterjee, M.: An Econometric Model for Resource Management in Competitive Wireless Data Networks. *IEEE Network* (November 2004)
- [2] Holma, H., Toscali, A.: *WCDMA for UMTS*. J. Wiley & Sons, Chichester (2001) Revised Edition
- [3] Perez-Romero, J., Sallent, O., Ruiz, D., Agusti, R.: An Admission Control Algorithm to Manage High Bit Rate Static Users in W-CDMA. In: 13th IST Mobile & Wireless Communications Summit, Lyon, France (June 2004)
- [4] Josephina, A., Andreas, P.: 4G Converged Environment: Modeling Network Selection as a Game. In: 16th IST Mobile and Wireless Communications Summit (2007)
- [5] Ormond, O., Murphy, J., Muntean, G.-M.: Utility-based Intelligent Network Selection in Beyond 3G Systems. In: 2006 IEEE International Conference on Communications, June 2006, vol. 4, pp. 1831–1836 (2006)
- [6] Niyato, D., Hossain, E.: Radio resource management games in wireless networks: an approach to bandwidth allocation and admission control for polling service in IEEE 802.16. *IEEE Wireless Communications* 14(1), 27–35 (2007)
- [7] Lin, H., et al.: ARC: An Integrated Admission and Rate Control Framework for Competitive Wireless CDMA Data Networks Using Noncooperative Games. *IEEE Trans. Mobile Comp.* 4(3), 243–258 (2005)
- [8] Vlacheas, P., Charilas, D., Tragos, E., Markaki, O.: Maximizing Quality of Service for Customers and Revenue for Service Providers through a Noncooperative Admission Control Game. In: *ICT Mobile Summit 2008*, Stockholm (June 2008)
- [9] Charilas, D., Markaki, O., Tragos, E.: A Theoretical Scheme for applying game theory and network selection mechanisms in access admission control. In: *International Symposium on Wireless Pervasive Computing (ISWPC)* (May 2008)
- [10] de Sousa Jr., V.A., de, R.A., Neto, O., de, F., Chaves, S., da Silva, A.P., Cavalcanti, F.R.P.: Conception and Evaluation of Access Selection Algorithms for Cooperative Beyond 3G Systems. In: *VI International Telecommunications Symposium (ITS 2006)*, Fortaleza-CE, Brazil, September 3-6 (2006)
- [11] Charilas, D., Markaki, O., Nikitopoulos, D., Theologou, M.: Packet-Switched Network Selection with the Highest QoS in 4G Networks. *Elsevier Computer Networks* 52(1), 248–258 (2008)
- [12] Markaki, O., Charilas, D., Nikitopoulos, D.: Enhancing Quality of Experience in Next Generation Networks through Network Selection Mechanisms. In: *Mobile Terminal Assisted Enhanced Services Provisioning in a B3G Environment Workshop, PIMRC* (September 2007)