A Fair and Dynamic Auction-based Resource Allocation Scheme for Wireless Mobile Networks

Tarik Taleb1,* and Abdelhamid Nafaa2
1Tohoku University, Japan, 2Dublin College University, Ireland
*talebtarik@ieee.org

Abstract—In this paper, we propose a fair and dynamic auction-based QoS negotiation scheme that allows users to dynamically negotiate their agreed service levels with their service provider. The scheme has three major design goals: ensuring a high degree of fairness among competing users, efficient utilization of the underlying network resources, and maximization of the service provider’s revenue.

A mathematical analysis is provided to demonstrate that when the three design goals are taken into account, the resource allocation function proposed in this paper represents a Pareto optimal solution.

I. INTRODUCTION

Generally speaking, Quality of Service (QoS) provisioning consists of two major operations: dynamic service level negotiation or management, and resource allocation. The former addresses the issue of QoS continuity when end users roam over different wireless networks, while the latter refers to the operation of enforcing the negotiated and agreed QoS terms. With this respect, the authors have recently proposed a scalable and prompt mechanism for dynamic Service Level Specification (SLS) negotiation in next generation wireless mobile networks [1].

In any communication system, the usefulness of dynamic SLS negotiation mechanisms hinges on an efficient resource allocation strategy. In resource allocation, a service provider finds optimal allocations of network resources to meet the service contract between a client and the service provider. A fundamental characteristic of wireless mobile environments is that demands for network resources are time-varying (due to the mobility of users). In such environments, the resource allocation should be dynamic and adaptive to changes in network conditions. Indeed, when the network is about to get congested, a service provider can offer some privileges to subscribers that accept to downgrade their current service levels. Similarly, if sufficient network resources become available, the service provider can encourage subscribers to join high service levels for better QoS. For this purpose, it is essential to develop a pricing scheme that prioritizes competition for resources among users, in other words, a strategy that allows mobile users to bid for network resources based on the competitiveness of their associated budgets.

In this paper, an auction-based admission control and bandwidth allocation policy is proposed. When demand exceeds supply (i.e., due to the arrival of new subscribers to a wireless network), the service provider runs an auction to determine the set of users that will be served and the corresponding service level of each. The resource allocation operation aims at achieving three major design goals: insurance of fairness among competing users, efficient utilization of network resources, and guarantee of the highest profit for the service provider.

The paper is organized in the following fashion. Section II highlights the relevance of this work to the state-of-art of auction-based resource allocation techniques. Section III describes the envisioned network architecture and states the resource allocation problem via a simple example. Section IV formulates the problem and analyzes the proposed auction-based resource allocation policy. The section also provides a mathematical model of the proposed solution. Section V presents a simple evaluation of the proposed scheme. The paper concludes in Section VI.

II. RELATED WORK

For resource allocation in wireless networks, a large body of research work has been done [4]. A large portion of these pioneering researches aims at achieving higher network utilization through an optimal allocation of resources. A highly missing point in most of these researches consists in the fact that they do not provide any motivation to encourage users to free unused resources when network resources become scarce due to excess in demand. Additionally, their adopted pricing models consider fix prices during the contract period.

As network resources are highly time varying in wireless mobile networks (due to the mobility of users), price-fixed models do not fairly reflect the varying market value of the network resources. To cope with this, in [5] network elements compute the price of network resources (e.g., bandwidth) as a function of the local supply and demand, and constantly inform brokers of the current price. Users then contact brokers and purchase resources from them. This concept, however, does not encourage competition among users and do not assist users to prevent service degradation when competition becomes high.

To address this issue, some researchers have investigated the idea of allowing users to bid against each other for network resources. Auction-based resource allocation schemes have been thus developed. In [6], a bandwidth broker is located at each sub-network and is running auction to allocate bandwidth among competing users based on their offered prices. In [7], a flexible auction based pricing scheme is proposed. In this scheme, the bandwidth-unit price a client is willing to pay is expressed as a function of time. This exempts users from periodically sending signaling messages to renegotiate their service requirements and gives more flexibility to network brokers in making their resource allocation decision. Most of the resource allocation schemes proposed in the recent literature aim at maximizing the data throughput without any

978-1-4244-2075-9/08/$25.00 ©2008 IEEE
consideration of fairness in service among competing users [8][9].

Other research works considered fairness but each with a different notion of fairness. In [10], max-min fairness in bandwidth allocation is achieved by allocating resources to unfortunate terminals (that exhibit the poorest performance) while maintaining a good utilization of the wireless network resources. In [11], a game theoretical approach is explored whereby the service provider attempts to maximize its revenue and the users attempt to maximize their individual utilities. In [12], an auction-based game theoretical approach is proposed. The scheme allows users to bid for a wireless channel using the second price auction mechanism.

III. PROBLEM STATEMENT

The envisioned architecture consists of a number of access points forming different domains, potentially administrated by different network operators (NOs). Each domain is administrated by a Global Service Negotiation Manager (GSNM) and an Authentication, Authorization, and Accounting (AAA) server. The GSNM server operates as a resource broker and carries out the service level negotiation procedure. At the GSNM server, different service levels are available. The GSNM server sets a minimum price and a maximum price for each service level (as a function of the varying channel quality conditions and offered QoS metric).

A user subscribed to a given service level will be charged for a price from within the corresponding minimum and maximum prices. It is assumed that each user possesses an initial amount of money. The mechanisms by which GSNMs admit or turn down requests, or allocate resources for users follow our proposed resource allocation strategy as will be explained later. GSNM allows users to compete for the wireless network resources. Naturally, users are interested in getting high throughput for the most reasonable price. On the other hand, GSNM is interested in maximizing the network operator’s revenue. This gives rise to an auction where users are given responsibility to determine their throughput.

In economic theory, there is a wide variety of algorithms. Notable examples are the all-pay auction, first-price auction, and second-price (or Vickrey) auction algorithms. In the all-pay auction algorithm, bidders independently submit single bids for an object. The object is sold to the bidder who makes the highest bid. However, the other bidders still have to pay their bid despite their failure in winning the auction. In the first price auction algorithm, the object is given to the bidder with the highest bid. Losers do not have to pay. In the second-price auction algorithm, the winner is intuitively the bidder with the highest bid. The object is however sold for a price equal to the second highest bid. In [12], the second price algorithm is used for resource allocation in wireless networks. It is demonstrated that it yields a good allocation of network resources. Whilst the proposed scheme can be implemented on top of any existing auction algorithm, we do consider the case of the first-price auction algorithm.

In the remainder of this section, we demonstrate via a simple example that current auction algorithms still fail in guaranteeing the best use of network resources, maximizing the revenue of a network operator while fairly satisfying the expectations of users. To illustrate the idea with more clarity, we consider the following scenario. We consider the case of a single GSNM with two subscribers A and B, each with an initial budget worth 7.5 and 8 money unit, respectively. We assume that the two users will be served for the same period of time. We also assume that the maximum bandwidth that can be served by the GSNM domain is 100kbps. We consider the case when GSNM provides four service levels as indicated in Table I. The minimum and maximum prices (in money unit) of each service level are listed in the table. For the sake of simplicity, we ignore the variations of channel conditions and assume the prices constant.

<table>
<thead>
<tr>
<th>Service level</th>
<th>Bandwidth (kbps)</th>
<th>Min. Price</th>
<th>Max. Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>25</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>L2</td>
<td>50</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>L3</td>
<td>75</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>L4</td>
<td>100</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

In this scenario, both users can afford service level L3. They can thus compete against each other for this service. We adopt a pricing scheme that has the following feature. If a user, with an initial budget that makes it eligible for service level L_i, gets its request downgraded to service level L_j (j < i), the user will be charged for the maximum price of service level L_j. This will make the user have the highest bid on service level L_j (on top of other users competing for service level L_j) and prevent it from experiencing further downgrades in its requested service level. Let S_{ik} denote the resource allocation strategy where users A and B are allocated service levels L_i and L_k, respectively (when j, k = 0, the request is rejected). Given the fact that user A bids an amount of money smaller than user B, the latter should be always allocated a service higher or similar to that of user A (j ≤ k). Table II lists all the possible resource allocation strategies along with the total required bandwidth, the total revenue, and fairness in the users’ satisfaction. The satisfaction of a user C (C ∈ {A, B}) is computed as follows:

\[ U(C) = \frac{bw_i}{bw_k} \cdot \frac{bw_i}{\alpha_{C,i}} \]  

(1)

where i and k denote the index of the allocated service level and the requested one, respectively. bw_i and bw_k denote the bandwidth provided by service levels L_i and L_k, respectively. \( \alpha_{C,i} \) is the actual price User C paid to subscribe to service level L_i. The rational behind this definition of the satisfaction metric is two fold: first to reflect how much a user gets its initially requested service level downgraded, second to indicate the bandwidth unit price at which the user paid for the service. The fairness index is computed as follows:

\[ F = \frac{(U(A) + U(B))^2}{2 \cdot (U(A)^2 + U(B)^2)} \]  

(2)

In the above scenario, the two strategies S_{23} and S_{31} cannot take place as the total required bandwidth exceeds the available
network resources 100kbps. There are two strategies that make full utilization of the network resources; \( S_{22} \) and \( S_{13} \). By applying a simple auction mechanism that merely allocates resources to the winning bidder and does not incorporate fairness (an equally important metric), the GSNM may allocate service level \( L_3 \) to user B as it makes the highest bid. To make full use of the network, user A will be then allocated to service level \( L_1 \). This strategy \( S_{13} \) will lead to a revenue of 11 and a fairness index equal to 0.772. However, by having an auction mechanism that can downgrade the service level of user B (winner when traditional auction algorithms are in use) to \( L_2 \) and allocating the same service level to user A, the network achieves its best performance and the network operator gets the maximum revenue. The system fairness also becomes perfect.

From the aforementioned example, it can be deduced that the use of traditional auction algorithms may favor users that make the highest bid and allocates them their requested service levels. However this comes at the cost of reduced revenue and poor fairness. A new auction-based resource allocation algorithm that takes into account system fairness, service provider’s revenue, and network utilization is required.

IV. PROPOSED AUCTION-BASED RESOURCE ALLOCATION SCHEME

In resource allocation, our focus is on bandwidth. The total bandwidth of the network is denoted as \( Bw \). We assume that \( N \) users are competing for the network bandwidth, each with an initial budget \( B_i \) \((i \in [1, N])\) and a call duration worth \( \theta_i \) time unit. The unit of the initial budget is defined as money unit per time unit. Without loss of generality, we assume that \((B_1 \leq B_2 \ldots \leq B_N)\).

GSNM is assumed to serve \( M \) service levels, \( L_j \) \((j \in [1, M])\). A user subscribing to service level \( L_k \) is allocated a portion of bandwidth equal to \( Bw_k \). Without loss of generality, we assume that the higher the index of a service level is, the higher its offered bandwidth is. Each service level \( L_1 \) has a low and upper bound prices, \( P_{1,\text{min}} \) and \( P_{1,\text{max}} \). To reflect the channel conditions in the service pricing, we assume that the prices of each service level are set proportionally to their offered bandwidth as follows:

\[
\frac{Bw_j}{P_{\text{max}, j}} = cst_2 \quad \forall j \in [1, M] \tag{3}
\]

Without loss of generality, we also assume that

\[
P_{1,\text{min}} < P_{1,\text{max}} \leq P_{2,\text{min}} < \ldots < P_{M,\text{min}} < P_{M,\text{max}} \tag{5}
\]

Let \( x_j \) denote the number of subscribers to service level \( L_j \). Let \( \alpha_{i,j} \) be the price user \( i \) actually pays for service level \( L_j \) per time unit. The following expresses money constraints.

\[
P_{j,\text{min}} \leq \alpha_{i,j} \leq P_{j,\text{max}} \tag{6}
\]

\[
\alpha_{i,j} \leq B_i \tag{7}
\]

If due to lack in network resources or tough competition, a user \( i \), eligible for service level \( L_j \), gets its requested service level downgraded to a service level \( L_k \) \((k \leq j)\), the user will be charged for the maximum price of \( L_k \), that is \( P_{k,\text{max}} \). From the user perspective, a user \( i \) is naturally always interested in subscribing to the highest possible service level with the most reasonable price. From the system perspective, it is desirable to maximize the revenue. In our auction strategy, we want to provide a fair system that makes the best use of the network resources, fairly satisfies the requests of all users, and maximizes the service revenue. This can be translated into the following equations.

\[
\text{Minimize} \left(N - \sum_{j=1}^{M} x_j\right) \tag{8}
\]

\[
\text{Minimize} \left(Bw - \sum_{j=1}^{M} Bw_j \cdot x_j\right) \tag{9}
\]

\[
\text{Maximize} \left(\theta_i \sum_{j=1}^{M} \alpha_{i,j}^*\right) \tag{10}
\]

where

\[
\alpha_{i,j}^* = \begin{cases} \alpha_{i,j} & \text{if user } i \text{ subscribes to } L_j \\ 0 & \text{otherwise} \end{cases} \tag{11}
\]

Note that in Equation 10, the call duration of each user is used. This is for the purpose of guaranteeing high revenue in the long run. Furthermore, while Equation 8 attempts to increase the scalability of the system by satisfying as many requests as possible, it does not guarantee a fair service to all competing users. To reflect system fairness, we consider the use of users’ satisfaction metric as defined below. Let user \( i \) requests subscription to service level \( L_j \) whereas it is allocated service level \( L_k \). The satisfaction of user \( i \) is measured as follows.

\[
U(i) = \frac{bw_k}{bw_j} \cdot \frac{bw_k}{\alpha_{i,k}} \tag{12}
\]

The fairness index is computed as follows.

\[
F = \frac{\left(\sum_{i=1}^{N} U(i)\right)^2}{N \cdot \left(\sum_{i=1}^{N} U(i)^2\right)} \tag{13}
\]

To guarantee fairness, our resource allocation strategy should maximize \( F \) in addition to satisfying Equations 8, 9, and 10.
Resource allocation in the envisioned architecture can be seen as a set of mixed strategies for finite, non-cooperative games between the mobile users. In the theory of non-cooperative games, this is known as Nash game. In the remainder of this section, we analytically demonstrate that when the three constraints (network utilization, fairness and revenue) are taken into consideration, our proposed resource allocation scheme provides a Pareto optimal solution, in other words unique Nash equilibrium.

First, we demonstrate that the resource allocation in case of $N$ ($3 \leq N$) users can be simplified to the case of two users. *Lemma 1*: If the proposed scheme can provide a unique and optimal solution in the case of two users, it can do the same for $N$ ($3 \leq N$) users.

*Proof*: The proof of this lemma can be done in a recursive manner with respect to $N$. Let $Bw$ denote the total available bandwidth that can be allocated to all the users. From the condition of the lemma, an optimal and unique allocation can be found for $N = 2$. Let assume that the lemma holds for up to the case of $(N-1)$ users, we prove that there is an optimal allocation in case of $N$ users. For $(N-1)$ users, each with an initial budget $B_i$, from the assumption, there is an optimal strategy, $S_{N-1}^*$, $(Bw, \{B_1, B_2, \ldots, B_{N-1}\})$, where the $Bw$ bandwidth of the network is optimally allocated. Now let assume that the $N$th user has an initial budget of $B_N$ and is eligible for service levels $L_k$ or lower. For the $N$ users, there are thus finite number ($k$) of strategies for bandwidth allocations:

$$S_N^* = \{S_{N-1}^*(Bw-Bw_j, (B_1, B_2, \ldots, B_{N-1})), Bw_j\}$$

where $(j \in [1, k])$.

Considering the $(N-1)$ users as a single user that has an initial budget worth $(\sum_{i=1}^{N-1} B_i)$ and is requesting a service level that provides a bandwidth equal to $(Bw - Bw_j)$, and using the condition of Lemma 1, an optimal and unique resource allocation strategy can be found for the $(N-1)$ users and the $N$th user separately, say $S_N^*$. Again using the recursive assumption, an optimal and unique allocation of the $(Bw - Bw_m)$ bandwidth can be found for the $(N-1)$ users.

*Lemma 2*: The proposed scheme can provide a unique and optimal solution (Pareto optimal) when two users are competing for the network resources.

*Proof*: We consider two users A and B, each with an initial budget equal to $B_1$ and $B_2$. They are eligible to service levels $L_k$ and $L_l$, respectively. Without loss of generality, we assume $(B_1 \leq B_2)$. Assuming that the service levels are ordered according to their index, $L_k$ should be thus higher than $L_l$; $(k \leq l)$. Let $Bw$ again denote the total available bandwidth that can be allocated to the two users.

- Case 1 ($Bw_k + Bw_l \leq Bw$): In this case, users are simply allocated the service levels $L_k$ and $L_l$ they are eligible for.
- Case 2 ($Bw_k + Bw_l > Bw$): In this case, the two users will be assigned two service levels $L_x$ and $L_y$ subject to $(x \leq y)$, $(x \leq k)$, and $(y \leq l)$. Here, two cases can be envisioned.

$\diamond x < k$ and $y < l$: In this case, the prices users A and B will pay are $P_{max,x}$ and $P_{max,y}$, respectively. To ensure high fairness, both users should exhibit almost the same satisfaction.

$$\frac{U(A)}{U(B)} = (1 \pm \epsilon) \Leftrightarrow \frac{Bw_x}{Bw_k} - \frac{Bw_x}{P_{max,x}} = (1 \pm \epsilon) \frac{Bw_y}{Bw_l} - \frac{Bw_y}{P_{max,y}} \quad (14)$$

where $\epsilon$ is negligible $(0 \leq \epsilon \ll 1)$. From Equation 3, we obtain

$$Bw_x = (1 + \epsilon) \frac{Bw_k}{Bw_l} - \frac{Bw_y}{P_{max,y}} \quad (15)$$

From maximizing the utilization of the network resources, we obtain

$$Bw_x + Bw_y = Bw \implies Bw_x = \frac{(1 + \epsilon) - 1}{\frac{1}{Bw_l} - \frac{1}{Bw_k}} \quad (16)$$

In this way, $x$ and $y$ are the index of the service levels whose bandwidths are closest to the values that can be obtained from Equations 15 and 16. It should be noted that since the price of service levels is proportional to the allocated bandwidth, the total revenue of the whole system can be maximized by maximizing the utilization of the network resources. It should be observed that from Equation 16, $Bw_y$ is unique. From Equation 15, the value of $Bw_x$ is also unique. It should be also remarked that a movement from the obtained allocation to a different one by modifying the values of $Bw_y$ or $Bw_x$ will affect the link utilization even if we guarantee high fairness, and vice versa. This shall make one user better off while the other user will be made worse off. This indicates the Pareto optimality of the obtained solution when the three objectives are taken into account. To conclude, the values of $x$ and $y$ represent a unique and optimal solution for both users A and B that satisfies the three objectives of our proposed scheme.

- $(x = k)$ or $(y = l)$: The values of $x$ and $y$ can be derived in the same manner as in the previous case. The only change will be in the price that will be paid by the users (e.g., in case of $(x = k)$, $\alpha_{A,k} = B_1$ and $\alpha_{B,y} = P_{max,y}$).

Using both Lemmas 1 and 2, we conclude that when the constraints on the system fairness, system revenue, and network utilization are taken into account, our proposed scheme provides a Pareto optimal resource allocation strategy to all competing users.

V. PERFORMANCE EVALUATION

While the performance of our proposed resource allocation mechanism can be evaluated considering the case of a large number of users and a general pricing scheme, for the sake of simplicity we consider the example provided in Section III.
with two users (1 and 2) and a simple pricing scheme as shown in Table I. We only vary the initial budgets of the two users ($B_1$ and $B_2$).

As previously discussed, in traditional auction-based resource allocation schemes, the user that makes the highest bid is allocated its requested service level. Other users get their requested service level downgraded if there is not much available bandwidth to satisfy their requests. In our proposed scheme, for the sake of a better fairness among competing users and higher revenue, even the user with the highest bid can get its requested service level downgraded. In the remainder of this section, we compare the performance of the proposed scheme against that of any traditional auction-based resource allocation mechanism. It should be noted that given the available $100\text{kbps}$ bandwidth in the considered example, our proposed scheme and traditional auction-based schemes will exhibit the same performance if the two users issue requests for service level $L_2$ or lower (Table I). We therefore consider the case when at least one end-user has an initial budget that makes it eligible for service level $L_3$ and beyond. Indeed, we consider the four following scenarios where the two users have initial budgets ($B_1$, $B_2$) equal to (5, 8), (7.5, 8), (7.5, 11), and (11, 11), respectively. According to Table I, in the four considered scenarios the two users are eligible for service levels ($L_2$, $L_3$), ($L_3$, $L_4$), ($L_3$, $L_4$), ($L_4$, $L_4$), respectively.

Fig. 1 plots the value of fairness index in case of both the proposed and traditional auction-based resource allocation schemes for the four considered scenarios. Results in terms of link utilization are omitted as in all considered scenarios, both schemes achieve 100% link utilization. As for the provider revenue, the proposed scheme always achieves the highest revenue ($=12$) compared to the traditional auction-based resource allocation scheme ($=11$). From Fig. 1, we observe that the proposed scheme achieves higher fairness. This is intuitively attributable to the features of our proposed scheme that downgrades the service level of even users with the highest bid, should that yield better fairness among users and higher revenue for the service provider.

VI. CONCLUSION

In this paper we analyzed the relationship between end-users and service providers from a service provisioning perspective. The focus was particularly on three important and inter-related aspects: (i) ensuring fairness among end-users, (ii) optimizing the usage of network resources, and (iii) maximizing the service provider’s revenue.

To reconcile these three design goals, we proposed an auction-based dynamic resource allocation scheme whereby the service provider sets a maximum and minimum price for each service class, and then runs a mathematical optimization model that finds out which end-users should be entitled to which services. The allocation scheme is based on a set of mixed strategies for finite, non-cooperative games between the mobile users (also known as Nash game theory). The proposed resource allocation scheme is dynamic in the sense that a user’s service level may be downgraded for the sake of increased fairness among users and/or higher revenue. Analytically, it is demonstrated that when the three design goals are taken into consideration, our resource allocation model reaches a "Nash Equilibrium" and converges to a Pareto optimal solution.

Future work will focus on integrating our proposed resource allocation scheme into a simulator to validate the mathematical analysis.

REFERENCES


