Utility-based optimal resource allocation in wireless networks

2 authors:

Wen-Hsing Kuo
Yuan Ze University
55 PUBLICATIONS **908** CITATIONS

SEE PROFILE

Wanjiun Liao
National Taiwan University
242 PUBLICATIONS **4,350** CITATIONS

SEE PROFILE

Utility-based Optimal Resource Allocation in Wireless Networks

Wen-Hsing Kuo and Wanjiun Liao

Department of Electrical Engineering
National Taiwan University
Taipei, Taiwan
Email: wjliao@ntu.edu.tw

Abstract- In this paper, we study utility-based optimal resource allocation in wireless networks. Utility here refers to a function which describes the degree of user satisfaction with a certain amount of allocated resource. We introduce utility function maximization in a centralized wireless environment and consider two types of traffic, i.e., best effort and QoS, the characteristics of which can be described by two existing utility functions. An allocation for the mixture of QoS and best effort traffics is proposed and its performance is evaluated via simulations. From this study, we show that optimal wireless resource allocation depends on the traffic types, the total available resource, and the channel quality, rather than solely dependent on the channel quality or traffic types as assumed in most existing work.

Keywords—utility optimization, resource allocation, wireless networks

I. INTRODUCTION

Resource allocation has been an active research topic in wireless networks [1-10, 12-17]. In such networks, radio resource is limited, and the channel quality of each user may vary. Given channel conditions and total available resource, system resource is allocated to users according to some performance metrics such as throughput and fairness [1-2] or according to the types of traffic [3]. "Throughput" and "fairness," however, are conflicting performance metrics. To maximize the system throughput, the system may allocate more resource to users with better channel conditions. This may cause radio resource monopolization by a small number of users, leading to unfairness. However, if the system attempts to provide fair treatment to all users, users with worse channel conditions tend to be allocated more resource so as to compensate for their channel conditions. Thus, the system throughput may be degraded dramatically. In [4-5], some mechanisms are proposed, which allow the system to behave either in "throughput-oriented" or "fairness-oriented" by adjusting certain parameters. However, they do not describe how to determine the approximate orientation, leaving this trade-off unsolved.

In this paper, we avoid such a "throughput-fairness" dilemma, and focus on "user satisfaction" for resource allocation. Since it is unlikely for the system to fully satisfy all users with different demands, the total degree of user satisfaction should be maximized. The degree of user satisfaction with a given amount of resource is described by a utility function. A utility function U(r) is a non-decreasing function with respect to allocated resource r. The more resource allocated, the more satisfied the user. The marginal utility function defined by $u(r) = \frac{dU(r)}{dr}$ is the derivative of the utility function U(r) with respect to the given resource r.

The exact expression of a utility function may depend on the traffic type, and can be obtained by studying the behavior and feeling of users. We leave the work of finding utility functions to psychologists and economists, and focus our work on maximizing the total utility with given utility functions.

We are not the pioneer to study utility-based resource management in wireless networks. Proposals [6-10] are examples of such work based on utility functions of different objectives for wireless networks. In [6], a power control scheme based on the utility function with respect to channel quality (i.e., signal-to-noise ratio, SIR) is proposed. In [7], a utility-based bandwidth allocation scheme is proposed, which adapts to the channel conditions and guarantees the minimum utility as requested by each user. In [8-9], a utility-based fairness allocation scheme is proposed. Resource is allocated to users so as to achieve an identical utility value for all users. However, letting users with different traffics to achieve an identical degree of satisfaction may not be a good way to use wireless resource efficiently. Worse, traffic difficult to satisfy tends to consume most system resource, causing another kind of unfairness. In [10], a utility-based scheduler together with a Forward Error Correction (FEC) and an ARQ scheme is proposed. This work, however, focuses mainly on the scheduling algorithm which gives lagging users more resource, leading to similar performance level for each user (i.e., by giving all users a fixed utility function). Besides, based on the considered utility functions, the proposed allocation is dependent on user channel conditions, and thus similar to the work proposed in [4].

Utility functions have been widely used in Internet pricing and congestion control [11]. Many bandwidth pricing schemes have been proposed for wireless networks [12-14]. The typical approach is to set the price for radio resource and allocate tokens to users so as to maximize the "social welfare" based on a users' bidding process. The bidding schemes, while useful for Internet pricing and congestion control, may not be practical for wireless networks. Since the traffic type, the number of users, and channel conditions are all time-varying in wireless environments, the bidding process is very expensive as the users must keep sending additional bidding data back and forth for real-time bidding. Besides, the control protocols of the wireless system must be modified to accommodate this process. Thus, wireless bidding is impractical to be implemented or integrated with wireless schedulers. Finally, the complexity and efficiency of wireless bidding has not been analyzed. This calls for optimal resource

allocation mechanisms without a wireless bidding process.

In this paper, we consider a wireless network with a central control system, such as the cellular base station, and maximize the total utility of all users for wireless resource allocation without a bidding process. Optimizing the total utility at the base station[1] can also maximize the "social welfare" as in a bidding scheme but in a simpler way. Instead of providing a specific scheduling algorithm, the ultimate goal of our research is to propose an optimal resource allocation algorithm under arbitrarily given utility functions. With this algorithm, we can help determine and validate the parameters[2] of existing schedulers. In this paper, we consider two common types of traffic: QoS and best effort traffics, the characteristic of which can be described by two utility functions [15]. Then we propose an allocation algorithm which allocates these two types of traffic according to their utilities. The allocation scheme is polynomial time solution and can be proved to be optimal under certain conditions. We also conduct simulations to verify the performance of the proposed scheme. From our study, we show that optimal wireless resource allocation depends on the traffic demand, total available resource, and wireless channel quality, rather than solely dependent on channel quality or traffic type as assumed in most existing work.

The rest of the paper is organized as follows. Sec. II describes the modeling of our system. Sec. III proposes a wireless allocation scheme and analyzes its performance. Sec. IV validates the performance of the proposed allocation schemes via simulation. Finally, the paper is concluded in Sec. V.

II. SYSTEM MODELING

Suppose that n users are served by a base station. Let r_{total} denote the total radio resource available at the base station, and r_i, the resource to be allocated to user i. Each user may have a different degree of satisfaction with a given resource, guided by the respective utility function of the traffic. Users with the same kind of traffic may not have the same utility function in a wireless network, because the wireless channel quality of each user may not be identical. Let q_i denote the channel quality of user i, $0 \leq q_i \leq 1$, and $i = 1, 2, \cdots n$. A smaller value of q_i indicates a worse channel quality. Given an amount of resource r_i and channel quality q_i, the amount of resource actually beneficial to user i is equal to $\theta_i = r_i \cdot q_i$. Therefore, the utility function of user i can be expressed as $U_i(r_i) = U(r_i q_i)$, where $U(.)$ is the utility

function of the traffic under consideration and $U_i(.)$ is the utility function for the type of traffic described by $U(.)$ but taking into account the channel quality of user i. The marginal utility function of $U(.)$ is $\dfrac{dU(r_i q_i)}{dr_i} = q_i \cdot u(r_i q_i)$, and that of $U_i(.)$ is $u_i(r_i)$.

The objective is to maximize $\displaystyle\sum_{i=1}^{n} U_i(r_i)$, subject to $\displaystyle\sum_{i=1}^{n} r_i \leq r_{total}$ and $\forall r_i \geq 0$.

Definition 2.1: A resource allocation $\Re^* = \{r_1, r_2, \cdots, r_n\}$ for n users is called an optimal allocation if for all feasible allocations $\Re_a = \{r_1', r_2', \cdots, r_n'\}$, $U(\Re^*) \geq U(\Re_a)$, where $U(\Re^*) = \displaystyle\sum_{i=1}^{n} U_i(r_i)$ and $U(\Re_a) = \displaystyle\sum_{j=1}^{n} U_j(r_j')$.

Note that the optimal allocation may not be unique in the system. To achieve optimal resource allocation in a wireless network, traffic of different utility functions should be treated differently. In this paper, we consider two types of traffic in wireless networks: best effort and QoS traffics. An empirical study [18] shows that the utility functions of "Hard QoS" (e.g., CBR) and elastic (i.e., best effort) traffics are a unit-step function and a concave function, respectively, with respect to allocated bandwidth, as depicted in Fig. 1. These two utility functions are defined as follows.

Definition 2.2: The unit-step utility function for QoS user i is described as $U_i(r) = U_{M_i} \times f_u(q_i r - r_{M_i})$, where $f_u(.)$ is a unit-step function, q_i is the channel quality of this user, M_i is the kind of QoS traffic for this user, r_{M_i} is the preferred allocated resource, and U_{M_i} is a normalization factor of this traffic, as shown in Fig. 1 (a).

Definition 2.3: A concave utility function $U(r)$ refers to a utility function with $u(r) > 0$ and $u'(r) < 0$ for all r, where $u(r) = \dfrac{dU(r)}{dr}$ and $u'(r) = \dfrac{du(r)}{dr}$.

By definition, a unit-step function is a discrete function, and a concave utility function must be a non-increasing and continuous function with respect to resource r. Fig 2 plots the marginal utility functions of the traffic shown in Fig. 1.
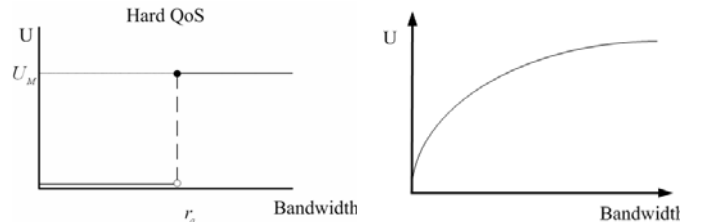
(a) Hard QoS traffic   (b) Best effort traffic

Figure 1. The utility functions of two types of traffic

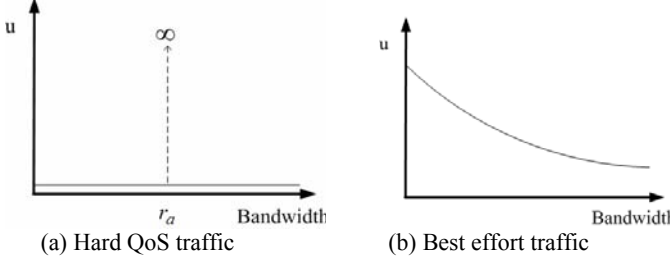(a) Hard QoS traffic   (b) Best effort traffic

Figure 2. The marginal utility functions of two types of traffic

III. PROPOSD SCHEME

To achieve optimal allocation, all available resource must be allocated so as to maximize the total utility. Thus, $\sum_{i=1}^{n} r_i = r_{total}$. Since best effort traffic has a concave utility function, to allocate best effort users optimally, it can be proved that all allocated best effort user (i.e. $r_i > 0$) must have an identical marginal value, $u_i(r_i)$. Due to space limitations, we will not include the proof in this paper. We next consider the allocation of QoS users. Since their utility function is discontinuous, they can only be allocated discretely, i.e., for all i, r_i is either $\frac{r_{M_i}}{q_i}$ or 0. We use $\frac{U_{M_i} q_i}{r_{M_i}}$ as the ith QoS user's average utility. All QoS users are served in decreasing order of their average utility.

Finally, we consider the co-existence of QoS and best effort traffics in the system. We define ΔU_i, the utility gain of the ith QoS user, as $\Delta U_i = U_{M_i} - \int_{r_{BE}-r_i}^{r_{BE}} u_{\Sigma}(r)dr$. If $\Delta U_i > 0$, then serving this QoS user increases the utility. To maximize total utility, the system must serve all QoS users with $\Delta U_i > 0$. Fig. 7 describes the proposed algorithm for a mixture of best effort and hard QoS traffics at a base station, where r_{BE} denotes the amount of residual resource given to best effort traffic and ΔU_i is the utility gain due to allocating resource r_i to QoS user i. The allocation rule of this mixed allocation is to: (i) allocate the first k best channel quality QoS users in the sorted queue, and (ii) then allocate the residual bandwidth (i.e., $r_{total} - r_{QoS}$) to all best effort users. The value of k is determined based on the requirements that there is sufficient resource for this QoS user (i.e., $r_{BE} - r_i > 0$) and the utility gain ΔU_i is positive.

This algorithm can be proved optimal if (a) each QoS user's traffic utility function is identical or (b) there are only best effort users in the system. Given QoS traffic with different utility functions, this allocation problem can be proved to be NP-complete and our proposed algorithm can also be proved to produce a solution which is a lower bound of the optimal solution. The worst-case complexity of this algorithm is O($m+n$), where m and n are the number of QoS users and best effort users, respectively.

Input: n users in the queue with channel qualities q_i and marginal utility functions $u_i(r)$, and r_{total}

Output: $\Re_{mix} = \{r_1, r_2, \cdots r_n\}$

Algorithm:

(1) Initialize $r_i = 0$, $i = 1,2,\cdots,n$ and $r_{BE} = r_{total}$

(2) Sort all QoS users i in descending order of $\frac{U_{M_i} q_i}{r_{M_i}}$, and store them in the queue

(3) Derive the total utility value of best effort users

Derive $u_j^{-1}(u)$ from $u_j(r)$ for all best effort users j

Derive $u_{\Sigma}^{-1}(u)$ for best effort users j, i.e.,

$$u_{\Sigma}^{-1}(u) = \sum_j u_j^{-1}(u)$$

Derive $u_{\Sigma}(u)$ from $u_{\Sigma}^{-1}(u)$

(4) If the queue is not empty, pop a QoS user i at the head of the queue. else, go to step (8)

(5) Allocate resource to the popped user i as follows:

$$r_i = \frac{r_{M_i}}{q_i}, \text{ if } r_{BE} \geq \frac{r_{M_i}}{q_i}$$

$r_i = 0$ and go back to step (4), otherwise

(6) Calculate utility gain of the ith QoS user ΔU_i, defined as

$$\Delta U_i = U_{M_i} - \int_{r_{BE}-r_i}^{r_{BE}} u_{\Sigma}(r)dr$$

(7) Check if the loop termination condition holds

If ($\Delta U_i > 0$), then

    $r_{BE} = r_{BE} - r_i$;

    go to step (4);

else

    $r_i = 0$;

    go to step (8);

(8) Determine r_j for all best effort users j based on the value of r_{BE}:

$$r_j = u_j^{-1}(u_{\Sigma}(r_{BE})), \text{ if } u_{\Sigma}(r_{BE}) < u_j(0)$$

$r_j = 0$, otherwise

Figure 7. Mixed allocation for both QoS and best effort traffic

IV. SIMULATION RESULT

In this section, we conduct simulations to verify our allocation algorithms. We consider pure QoS traffic, pure best effort traffic, and a mixture of both in the proposed algorithm. The simulation parameters are described as follows. For QoS traffic, the utility function is a unit-step function with $r_a = 10$ and $U_M = 1$, i.e., $U_{QoS}(r) = f_u(r-10)$; for best effort traffic, $U_{BE}(r) = 1 - e^{r/10}$. The value of q_i is randomly generated over [0, 1]. We then measure the distributions of r_i and θ_i under different values of r_{total}.

Fig. 9 shows the values of r_i and θ_i for QoS traffic. Fig. 9 (a) depicts that when r_{total} is small, the system allocates more resource to users with better channel

conditions; but as r_{total} increases, the allocated resource to users is still fixed either at r_a or 0, because the utility function is a unit-step function. Fig. 9 (b) shows that the actual resource obtained by each user, if allocated, is identical, i.e., at their θ_i. Fig. 10 shows the values of r_i and θ_i for best effort traffic. Fig. 10 (a) depicts that when r_{total} is small, the system allocates more resource to users with better channel conditions. Since the utility function of best-effort traffic satisfies $u(r) > 0$ for all r, all users demand resource all the time. Thus, as the value of r_{total} increases, the r_i values of all users increase. Fig. 10 (b) shows that users with larger q_i always have larger θ_i. Fig. 11 shows the values of r_i and θ_i for mixed traffic. Figs. 11 (a) and 11 (b) show the distributions of r_i and θ_i. Their characteristics are similar to those in Figs. 9 and 10.
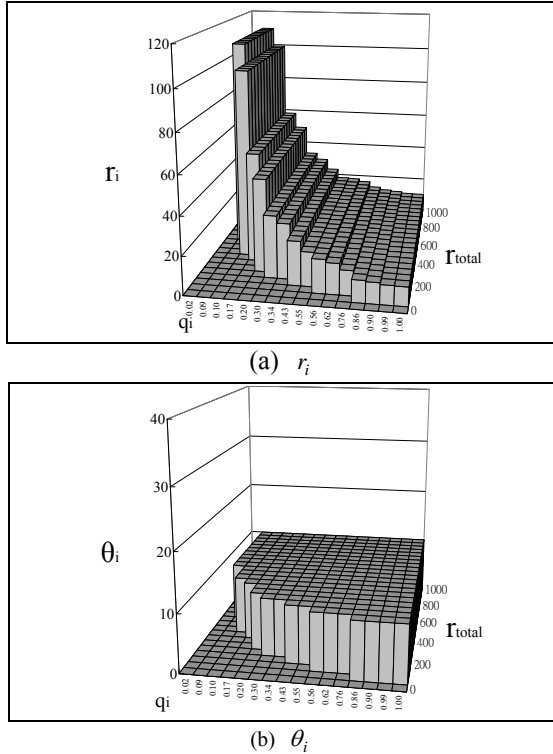

(a) r_i

(b) θ_i

Figure 9. Resource allocation for pure Hard QoS traffic
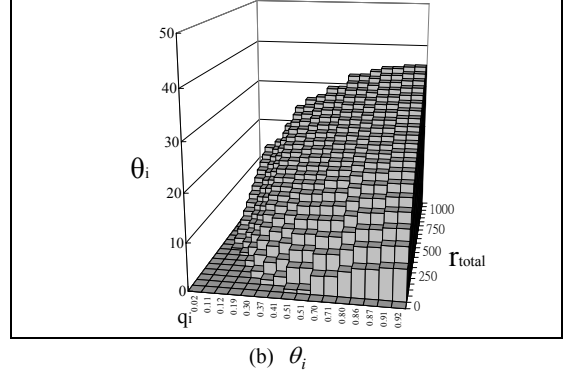

(a) r_i

(b) θ_i
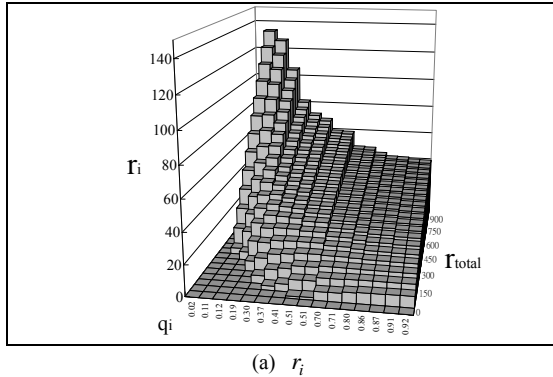
Figure 10. Resource allocation for pure best effort traffic
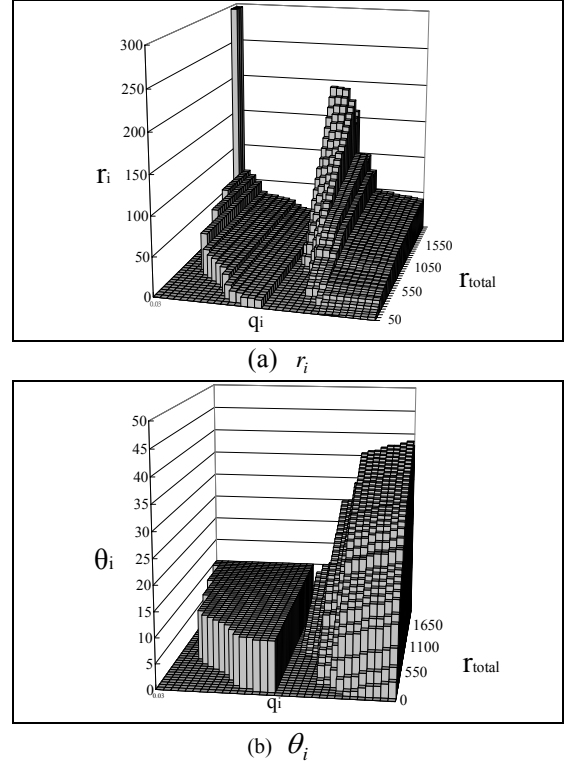

(a) r_i

(b) θ_i

Figure 11. Resource allocation for a mixture of QoS and best effort traffic

In Fig. 12, different resource allocation schemes are compared with the proposed allocation schemes. In [4], a scheme is proposed to allocate radio resource proportionally based on factor q_i^{α}. Depending on the setting of the value α, the system can be tuned to work for different performance metrics. The curve denoted "throughput" is for $\alpha = 1$, which gives more resources to users with better channel qualities, thereby leading to a larger system throughput. The curve denoted "fairness" is for $\alpha = -1$, giving all users an identical value of $\theta_i = r_i \cdot q_i$. The curve denoted "fixed" is for $\alpha = 0$, which provides the same amount of resource for all users. Note that the schemes proposed in [8-9] are examples of the "fairness" scheme (i.e., $\alpha = -1$), and the GR+ scheme in [1], an example of the "throughput" scheme.

4
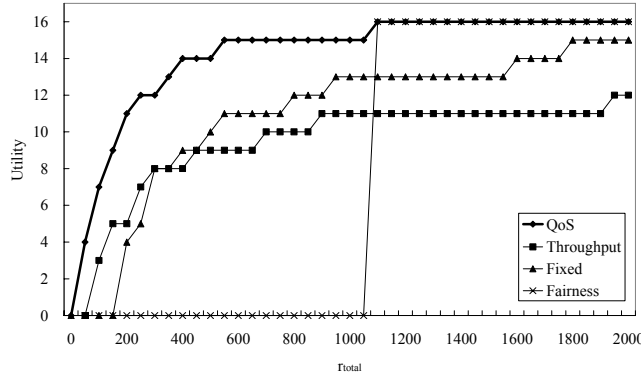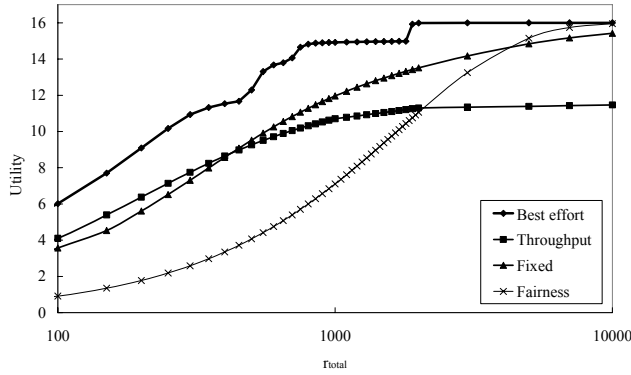
(a) QoS traffic

(b) Best effort traffic

Figure 12. Utility comparison with different resource allocation schemes

Fig. 12 (a) compares the total utility of QoS traffic for different allocation schemes, and Fig. 12 (b) gives that of best effort traffic. The QoS curve in Fig. 12 (a) and the "best effort" curve in Fig. 12 (b) are from the proposed mechanism. The results show that the proposed mechanism outperforms the other mechanisms. Of the other three mechanisms, the "throughput-first" scheme has a higher total utility when r_{total} is small, but the "fixed" allocation one is closer to the proposed one as r_{total} increases. Finally, when r_{total} becomes very large, the "fairness-first" scheme achieves a better utility. Note that the axis of r_{total} in Fig. 12 (b) is in the logarithmic scale.

V. CONCLUSION

In this paper, we study utility maximization for resource allocation in infrastructure-based wireless networks. We propose an algorithm which runs in polynomial time for mixture of QoS and best effort traffic. From the simulation results, we find that different types of traffic need different kinds of schemes to achieve optimal allocation. Also, when r_{total} is small, the system tends to allocate more resources to users with better channel conditions, i.e., "throughput-oriented;" however, when r_{total} is abundant, the system becomes "fairness-oriented," meaning that even with the same traffic, the preference orientation between throughput and fairness in order to achieve optimality can still differ. This leads us to conclude that existing channel-dependent-only resource schemes and schedulers cannot provide optimal allocation in wireless networks. An optimal radio resource allocation must depend on the traffic type, the total available resource r_{total} and the user channel quality q_i. To the best of our knowledge, this is the first work to address all three issues.

In this paper, we focus only on resource allocation and propose three polynomial time algorithms for different types of traffic in wireless networks. In the future, we will investigate utility-based call-admission control, and integrate it with the schemes proposed in this paper.

REFERENCES

[1] D. Angelini and M. Zorzi, "On the Throughput and Fairness Performance of Heterogeneous Downlink Packet Traffic in a Locally Centralized CDMA/TDD System," IEEE VTC 2002-Fall.

[2] A. Furuskar et al., "Performance of WCDMA High Speed Packet Data," IEEE VTC 2002-Spring.

[3] Cao, Y., Li, V.O.K., and Cao, Z., "Scheduling delay-sensitive and best-effort traffics in wireless networks," IEEE ICC, Anchorage, Alaska, May 2003, pp. 2208 – 2212.

[4] Z. Jiang, and N.K. Shankaranarayana, "Channel Quality Dependent Scheduling for Flexible Wireless Resource Management," IEEE Globecom '01.

[5] W.-T. Chen, K.-C. Shih, and J.-L. Chiang, "Flexible Packet Scheduling for Quality of Service Provisioning in Wireless Networks," Parallel and Distributed Systems, Dec. 2002.

[6] M. Xiao, N. B. Shroff and E. K. P. Chong "A Utility-Based Power-Control Scheme in Wireless Cellular Systems," IEEE/ACM Transactions on Networking, Vol. 11, No. 2, 2003.

[7] Y. Cao and V. O. K. Li, "Utility-Oriented Adaptive QoS and Bandwidth Allocation in Wireless Networks", IEEE ICC 2002, no. 1, April 2002, pp. 3071-3075.

[8] G. Bianchi and A.T. Campbell, "A Programmable MAC Framework for Utility-Based Adaptive Quality of Service Support," IEEE Journal on Selected Areas in Comm., 18(2): 244-255 (2000)

[9] R.-F. Liao and A. T. Campbell, "A Utility-Based Approach to Quantitative Adaptation in Wireless Packet Networks," ACM WINET, Vol. 7, No. 5, pp. 541-557, September 2001

[10] X. Gao, T. Nandagopal and V. Bharghavan, "Achieving Application Level Fairness Through Utility-based Wireless Fair Scheduling," IEEE Globecom 2001.

[11] F. P. Kelly, "Charging and Rate Control for Elastic Traffic" European Transaction on Telecommunication, January 1997.

[12] V. A. Siris, B. Briscoe and D. Songhurst, "Economic Models for Resource Control in Wireless Networks," IEEE PIMRC 2002, LISBON, PORTUGAL, SEP. 15-18, 2002.

[13] P. Marbach and R. Berry. "Downlink Resource Allocation and Pricing for Wireless Networks". IEEE INFOCOM 2002.

[14] P. Liu, R. Berry, M. L. Honig, and S. Jordan, "Slow-Rate Utility-Based Resource Allocation in Wireless Networks", IEEE GLOBECOM 2002, Nov. 2002.

[15] S. Shenker, "Fundamental Design Issues for the Future Internet," IEEE Journal on Selected Areas in Communications, 13(7): 1176-1188 (1995).