

Optimal Downlink Resource Allocation for Non-realtime Traffic in Cellular CDMA/TDMA Networks

Keivan Navaie, *Member, IEEE*, and Halim Yanikomeroglu, *Member, IEEE*

Abstract—In this letter, using a utility-based approach we formulate the optimal downlink resource allocation problem for non-realtime traffic in a CDMA/TDMA cellular network. We then introduce the notion of *multi-access-point diversity* which is a potential form of diversity in cellular networks, where a signal can be transmitted to the corresponding mobile user via multiple base-stations. A joint base-station assignment and packet scheduling method is then proposed which exploits multi-access-point and multi-user diversity in one shot. The simulation results of the proposed method show significant throughput improvement due to multi-access-point diversity gain.

Index Terms—Base-station assignment, CDMA/TDMA, diversity, packet scheduling, utility function.

I. INTRODUCTION

IN wireless cellular communications, delay tolerance of non-real time traffic can be exploited in transmission scheduling to improve system throughput (see e.g. [1]– [2]). In such techniques, at each time instant, the base-station (BS) transmits with its maximum power to a user when its time varying channel capacity happens to be at (or near) its peak while other users are kept inactive. The resulting throughput improvement is referred to as a *multi-user diversity* gain. In these methods, it is assumed that another functionality assigns a BS to each user as a server.

In order to facilitate the BS assignment, each BS broadcasts a pilot channel. In cellular networks, especially in dense areas with small size cells, a large number of users are able to receive pilot channels with acceptable quality from more than one BS. The set of these BSs is called the “active set.” The independent time variations of the channels between a user and the BSs in its active set introduce a type of diversity which we refer to as *multi-access-point diversity*.

In this letter we utilize the utility-based framework in [3] and formulate the optimal downlink resource allocation for a cellular network that supports non-realtime traffic. For each packet transmitted by a BS, we consider a utility function which depends on the corresponding channel bit-rate and the experienced delay. The total utility (the sum of the utilities of

all packets) is then considered as a total network performance indicator. We then show that maximizing the total utility results in joint BS assignment and packet scheduling; this approach jointly exploits multi-user and multi-access-point diversities. We further demonstrate that this problem can be mapped into a multi-dimensional multiple-choice knapsack problem (MMKP) that is NP-hard. A heuristic algorithm, presented in [4] with polynomial time computational complexity, is then adopted to solve this problem. A similar formulation has been used in our previous works [5]– [6] for downlink resource allocation in pure CDMA networks.

Joint BS assignment and resource allocation for downlink is considered in [7]– [8] as well. However, in both of these methods, separate BS assignment and resource allocation algorithms are proposed. These algorithms are then combined in a sequential manner for more efficient resource utilization. The distinguishing feature of our proposed method, compared to the previous works in [7]– [8], is that it gives in one shot, both the packets scheduled for transmission and the corresponding BSs to transmitting those packets. Moreover, since the BS assignment is performed in the same time-scale of packet scheduling, the proposed method will be able to exploit both multi-user and multi-access-point diversities. Furthermore, by adopting a heuristic polynomial time algorithm, it can be implemented with an acceptable level of complexity.

II. SYSTEM MODEL

We consider a CDMA/TDMA cellular network with time domain scheduling that supports only non-realtime traffic that can tolerate packet transmission delay. The system is time-slotted with each slot containing M frames of T_f seconds. In the time domain scheduling, the total transmit power of a BS is allocated to a single user within each frame while the rest of the users are kept inactive [1]– [2]. We also assume that the length of the time-slot is chosen so that the channel variations are negligible within the slot. Fixed L -bit packets are transmitted in an integer number of frames. At time t , there are $N(t)$ packets waiting in the system to be served. $\mathcal{B} = \{BS_1, \dots, BS_B\}$ is the set of B BSs in the network.

An active set $\mathcal{A}_i(t)$, corresponding to each packet i , is defined as the set of BSs which can be assigned to the destination user of packet i , $d(i)$, as the server:

$$\mathcal{A}_i(t) = \{j | j \in \mathcal{B}, \gamma_{d(i)j}(t) \geq \gamma_{min}\}, \quad (1)$$

In (1), $\gamma_{d(i)j}(t)$ is the E_c/I_0 level of the pilot channel of BS_j received by $d(i)$, and γ_{min} is the minimum required E_c/I_0 .

Manuscript received November 15, 2005. The associate editor coordinating the review of this letter and approving it for publication was Prof. Yu-Dong Yao. This work was supported by the Special Research Opportunities grant from Natural Sciences and Engineering Research Council (NSERC) for participation in the Wireless World Initiative New Radio (WINNER) project.

The authors are with the Broadband Communications and Wireless Systems (BCWS) Centre, Department of Systems and Computer Engineering, Carleton University, Ottawa, Ontario, Canada, K1S 5B6 (e-mail: {keivan,halim}@scc.carleton.ca).

Digital Object Identifier 10.1109/LCOMM.2006.04027.

$N_{\mathcal{A}_i} \leq N_{\mathcal{A}}$ indicates the number of BSs in $\mathcal{A}_i(t)$, where $N_{\mathcal{A}}$ is a system parameter. If $N_{\mathcal{A}_i(t)} > 1$, then packet i can potentially be transmitted by more than one BS.

Let $m_{ij}(t)$ be the number of required frames for transmission of packet i by BS_j at time t ,

$$m_{ij}(t) = \left\lceil \frac{L}{r_{ij}(t)T_f} \right\rceil, \quad (2)$$

where $\lceil \cdot \rceil$ gives the upper nearest integer since an integer number of frames should be allocated to each packet, and $r_{ij}(t)$ is the bit-rate of the channel between BS_j and $d(i)$ at time t . Note that $r_{ij}(t)$ is a function of different parameters including channel gain, BS transmit power, background noise, and inter-cell interference. We assume that the value of $r_{ij}(t)$ is sent to the network by the user as it is the case in 1xEV-DV [9]. Variable channel bit-rates are implemented by using the orthogonal variable spreading factor technique [10].

For a packet i , we associate a utility-function, $u_{ij}(t)$, that indicates the ‘‘profit’’ earned by the network as a result of transmitting packet i from BS j at time t . Utility function, $u_{ij}(t)$, serves as an optimization objective for packet transmission and is a function of the allocated network resources to that packet as well as the quality of service experienced by that packet. The earned profit modelled by the utility function provides a priority metric for each packet served by a given BS; the higher the value of a utility function, the higher the priority of transmitting the corresponding packet.

In the case of non-realtime traffic, the utility function is a function of $r_{ij}(t)$ and $\tau_i(t)$, where $\tau_i(t)$ denotes the amount of time that packet has spent in the system until time t . The profit earned by the network by serving a packet is an increasing function of that packet’s destined user’s wireless link quality. In addition, for two users with the same channel quality at time t , it would be more beneficial, from network point of view, to send the packet which experiences a higher delay. As long as having the general characteristics mentioned above, the formulation presented in this letter is independent from the exact definition of the utility function. Generally, utility function is defined by the service provider to quantify the trade-off between the performance (e.g., total throughput) and the fairness (see for instance, [11]).

III. DOWNLINK RESOURCE ALLOCATION PROBLEM

We define the total network utility at time t , $U(t)$, as the network performance indicator:

$$U(t) \triangleq \sum_{i=1}^{N(t)} \sum_{j=0}^B u_{ij}(t)b_{ij}(t), \quad (3)$$

where $b_{ij}(t)$ is the assignment indicator at time t ; $b_{ij}(t) = 1$ if packet i is transmitted by BS_j , and $b_{ij}(t) = 0$, otherwise. We also consider a null BS BS_0 with M virtual frames for which $u_{i0}(t) = 0$ for all packets. If $b_{i0}(t) = 1$, packet i is not scheduled for transmission at time t , therefore, $\tau_i(t+1) = \tau_i(t) + M.T_f$. We also define

$$\mathcal{B}_b(t) \triangleq \{b_{ij}(t) | i = 1, \dots, N(t), j = 0, \dots, B\}.$$

Our objective is to maximize the total network utility:

$$\max_{\mathcal{B}_b(t)} \sum_{i=1}^{N(t)} \sum_{j=0}^B u_{ij}(t)b_{ij}(t) \quad (4)$$

$$\text{s.t.} \sum_{i=1}^{N(t)} m_{ij}(t)b_{ij}(t) \leq M, \quad j = 1, \dots, B, \quad (5)$$

$$\sum_{j=0}^B b_{ij}(t) = 1, \quad i = 1, \dots, N(t), \quad (6)$$

$$b_{ij}(t) \in \{0, 1\}. \quad (7)$$

The inequality in (5) shows the downlink resource constraint for the BSs in \mathcal{B} . Furthermore, (6) demonstrates that each packet should be transmitted by only one BS. At each time t , the solution of (4) gives the packets that are scheduled for transmission. For each packet, it also gives the corresponding BS that is assigned as the server. This information can be communicated to the corresponding users via the signalling channels in the beginning of each time-slot.

For the specific case of a given BS assignment, the optimization problem in (4) results in a scheduling discipline in which, for each time-slot, packets are transmitted in the order of the value of their corresponding utility functions. In this case, the resulting scheduling only exploits the multi-user diversity. Moreover, in this case (4) results in proportional fairness scheduling if the utility function in [2] is used.

Since this optimization problem is solved for each time-slot, time index t is dropped for brevity. Here, we first define MMKP and subsequently we show that (4) is an MMKP.

MMKP [4]: Consider a knapsack with B distinct resources represented by (M_1, M_2, \dots, M_B) . There are N groups, each with K_i items. Each item j of a group i has a particular value, u_{ij} , and requires B distinct resources represented by the vector $(m_{ij1}, m_{ij2}, \dots, m_{ijB})$. The MMKP objective is to select one item from each group to maximize the total value of the collected items subject to B resource constraints.

Consider the cellular network as a knapsack with $B+1$ resource constraints. Each resource corresponds to a BS in the network coverage area in which we also include the virtual null BS. Resource constraints are the number of available frames in each BS in each time-slot (i.e., M). Each packet i as a group is associated with $K_i = N_{\mathcal{A}_i} + 1$ BSs as items, corresponding to the BSs in \mathcal{A}_i and the null BS. The value of the corresponding item is u_{ij} . For each item the required resource is $m_{ijk} = m_{ij}$ for $j = k$, and 0, otherwise. Therefore, the downlink resource allocation in (4) is an MMKP.

In each time-slot, each user obtains and sends the channel bit-rate to the BSs with acceptable E_c/I_0 ; therefore, for all packets, the network knows r_{ij} , for $i = 1, \dots, N$, $j \in \mathcal{A}_i$. Then the corresponding u_{ij} and m_{ij} are obtained.

Note that MMKP is NP-Hard [4]. Therefore, the computational complexity of finding the exact solution of (4) is in exponential order of the problem size. Obviously such solutions are not suitable for practical purposes. There are heuristics algorithms for MMKP with polynomial time complexity. As an example, in this letter we adopt a heuristic algorithm presented in [4] with a polynomial time complexity.

In practice, the cellular network can be partitioned into smaller parts; each part includes a number of BSs controlled by a radio network controller (RNC). In cases where some of the BSs in the active set of a packet are controlled by different RNCs, signalling between the RNCs is required.

IV. SIMULATION RESULTS

The simulated cellular network is a two-tier hexagonal cell configuration including 19 cells ($B = 19$), based on UMTS standard; the wrap-around technique is also employed. No adaptive modulation is considered. The number of active users in the network is N_u . The simulation parameters are presented in Table I. The pilot channel powers are adjusted so that 40% of the users receive the pilot channel of more than one BS with acceptable quality. Non-uniform user distribution is expressed by the non-uniformity factor, μ_D so that $(1 - \mu_D)N_u$ users are distributed uniformly and the rest of the users are distributed within randomly located hot-spots.

For each user, fixed length packets are generated by a Poisson arrival process with an average rate of λ packets per second. For a packet i transmitted by BS_j in time-slot t the following utility function, $u_{ij}(t)$, $j \neq 0$ is defined:

$$u_{ij}(t) \triangleq \frac{1}{m_{ij}(t)} \exp(\tau_i(t) - \bar{\tau}(t)), \quad (8)$$

where $\bar{\tau}(t) = \frac{1}{N(t)} \sum_i \tau_i(t)$ is the average delay. Note that in (8), a packet is given a large utility either when the corresponding user experiences a “good” channel condition or when it experiences a “bad” relative delay status. Parameter $\tau_i(t)$ is available in the BS and $m_{ij}(t)$ can be obtained based on $r_{ij}(t)$ s sent by the users to the BSs through (2), which is a function of the wireless channel condition between the user $d(i)$ and BS j . The utility function in (8) is similar to that given in [11] which attempts fairness in delay.

In order to study the multi-access-point diversity gain, three different systems are considered. In System I, for each user, the BS with maximum E_c/I_0 level is assigned as the server and packets are transmitted in first-come-first-serve fashion using a time domain scheduling scheme. In System II, the BS assignment is similar to that in System I, however, packets are scheduled so that the total utility of each cell is maximized; i.e., only multi-user diversity is exploited. System III uses our proposed joint BS assignment and packet scheduling.

We normalize the average achieved throughput of Systems II and III by the average achieved throughput of System I. Fig. 1 illustrates the normalized average achieved throughput versus the average number of users in each cell, N_u/B . Two different user spatial distributions are considered: uniform ($\mu_D = 0$) and non-uniform ($\mu_D = 0.4$). Note that the throughput gain of System II is due to the multi-user diversity alone. The difference between the throughput gains of System II and III indicates the multi-access-point diversity gain. Furthermore, as demonstrated, the multi-access-point diversity gain is increased by the non-uniformity of the users’ spatial distribution.

We also compare the average throughput of the proposed method in this letter with the one in our previous works in [5]–[6] with $N_u/B = 20$. An average throughput improvement of 21.17% (27.63%) is observed for $\mu_D = 0$ ($\mu_D = 0.4$).

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Cell radius	1000 m
BSs transmit power	10 W
N_A	4
Standard deviation of log-normal fading	8 dB
Thermal noise density	-174.0 dBm/Hz
Propagation loss exponent	4
Frame length (T_f)	10 ms
Number of frames in each time-slot (M)	10
Available bit-rates r_{ij}	16, 32, 64, 144, 384 kb/s
Required E_c/I_0 (γ_{min})	-13 dB
Packet arrival rate (λ)	10 packets/s

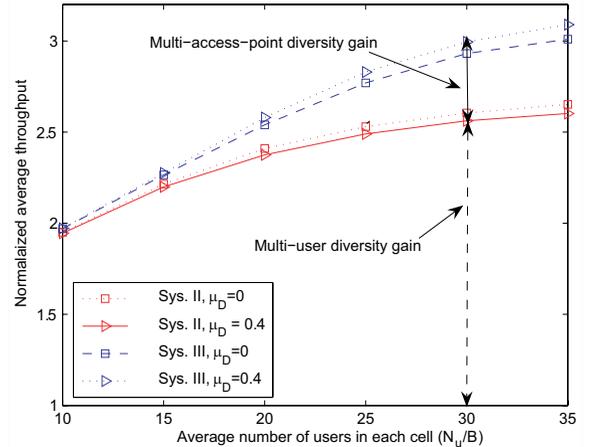


Fig. 1. Normalized average achieved network throughput of Systems II and III versus average number of users in each cell for $\mu_D = 0$ and $\mu_D = 0.4$.

The achieved gain is due to the exploitation of the multi-user diversity and multi-access-point diversity.

V. CONCLUSION

In this letter, we propose a joint BS assignment and packet scheduling method that exploits, in one shot, both multi-user and multi-access-point diversities. The simulation results show that using the proposed method substantially improves the system performance.

REFERENCES

- [1] P. Bender *et al.*, “CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users,” *IEEE Commun. Mag.*, vol. 38, pp. 70–77, July 2000.
- [2] P. Viswanath *et al.*, “Opportunistic beam-forming using dumb antennas,” *IEEE Trans. Info. Theory*, vol. 48, pp. 1277–1294, June 2002.
- [3] F. P. Kelly, “Charging and rate control for elastic traffic,” *European Trans. Telecommun.*, vol. 8, pp. 33–37, Jan. 1997.
- [4] M. Moser *et al.*, “An algorithm for the multidimensional multiple-choice knapsack problem,” *IEICE Trans. Fundamentals*, vol. E80-A, pp. 582–589, Mar. 1997.
- [5] M. Shabani and K. Navaie, “Joint pilot power adjustment and base station assignment for data traffic in cellular CDMA networks,” in *Proc. IEEE/Sarnoff Symp. 2004*, pp. 179–1831.
- [6] M. Shabani *et al.*, “Downlink resource allocation for data traffic in heterogeneous cellular CDMA networks,” in *Proc. IEEE ISCC’04*, vol. 1, pp. 436–441.
- [7] J. W. Lee *et al.*, “Joint resource allocation and base-station assignment for the downlink in CDMA networks,” preprint, 2003.
- [8] C. Makaya and S. Aissa, “Joint scheduling and base station assignment for CDMA packet data networks,” in *Proc. IEEE VTC 2003-Fall*, vol. 3, pp. 1693–1697.
- [9] 3GPP2, “Introduction to 1xEV-DV air interface,” 2002.
- [10] H. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. New York: John Wiley and Sons, 2000.
- [11] S. Shakkottai and A. Stolyar, “Scheduling for multiple flows sharing a time-varying channel: the exponential rule,” *American Mathematical Society Translations*, ser. 2, Y. M. Suhov, Ed., 2002.