Incentive-compatible Resource Allocation in Overlapping Heterogeneous Wireless Networks

Bin Bin Chen

Department of Computer Science, National University of Singapore Email: chenbinb@comp.nus.edu.sg

Mun Choon Chan

Department of Computer Science, National University of Singapore Email: chanmc@comp.nus.edu.sg

Abstract

This paper considers the *coordinated radio resource allocation problem* for users which are simultaneously covered by multiple overlapping heterogeneous wireless networks. As the resource allocation decision depends on the channel measurement and feedback from users, inefficiency and instability arise if a selfish user can manipulate its measured channel state to increase its gain from network. Our contribution in this paper is the introduction of incentive compatibility as an addition criterion in the design of a resource allocation scheme.

We formulate the *multi-cell resource allocation game* to capture the strategic interactions among users. A resource allocation scheme is incentive compatible if each user's dominant strategy under the resulted game is to honestly report its channel state. We consider both multi-association setting, where a MS is allowed to simultaneously associate with multiple BSs, and single-association setting, where a MS is only associated with one BS. We show that for multi-association setting, a natural generalization of proportional fair allocation is incentive compatible. In contrast, the optimal solution using the same fairness criterion under single-association is not incentive compatible. In order to exploit the benefit of single-association, we propose an allocation scheme based on selfish load balancing. We show that such a scheme always converges to a Nash equilibrium, and achieves performance close to the optimal single-association allocation.

I. INTRODUCTION

Overlapping coverage of wireless base stations (BSs)¹ is a common phenomenon in mobile communication systems. For a particular radio access network, neighboring cells or sectors overlap with each other. In addition, deployment and inter-operation of a wide array of wireless access networks, ranging from 3G network to Wi-Fi hotspots, open the opportunity of overlapping coverage from BSs using heterogeneous radio access technologies. In

¹We use BS as a general term to refer to both 3G base station and Wi-Fi access point (AP).

such an environment, a multi-mode mobile station (MS) can flexibly associate with one or simultaneously multiple BSs.

In an overlapping multi-cell heterogeneous wireless networks, the coordinated resource allocation decision can be decomposed to two layers: the *inter-cell association control layer* that decides which BS(s) a MS should associate with, and the *intra-cell allocation layer* that determines how radio resource of a single BS should be shared among its associated MSs. The traditional criteria to judge a resource allocation scheme include efficiency, fairness and load balancing.

Proportional fair allocation [1] is widely accepted as an appropriate allocation scheme for elastic traffic in wireless networks [2] [3], as it strikes a good balance between efficiency and fairness. The tradeoff is particularly important when MSs use different data rate to communicate with the same BS according to their various channel conditions. The concept has been generalized to multi-cell environment [4] [5] [6], thus fairness is considered in a global sense, with load-balancing among cells naturally incorporated into the definition.

However, implementing multi-cell proportional fair allocation requires the channel state information of adjacent MS-BS pairs to be known. In practice, the channel state for transmission from BS to MS is measured by individual MS, which then periodically feeds the channel state back to the resource allocator. As a result, it is possible for an intelligent and selfish MS to manipulate the reported channel states to increase its own resource allocation, while causing problem of inefficiency and instability for the system. Our contribution in this paper is the introduction of *incentive compatibility* as an additional criterion in multi-cell resource allocation.

We formulate the *multi-cell resource allocation game* to capture the selfish behavior of users. The game defined by a given resource allocation scheme is said to be *incentive compatible*, if the dominant strategy for each player is to honestly measure and report its actual channel state.

We consider both multi-association setting, where a MS is allowed to simultaneously associate with multiple BSs, and single-association setting, where a MS is only associated with one BS. Our result shows that for multiassociation setting, a natural generalization of proportional fair allocation (Coordinated Proportional Fairness or CPF), which can be efficiently solved as a convex programming problem, is incentive compatible. In contrast, the optimal solution using the same fairness criterion under single-association is not incentive compatible. In order to exploit the benefit of single-association, we propose an allocation scheme based on selfish load balancing (SLB). We show that SLB always converges to a Nash equilibrium. Evaluation results show that SLB converges quickly and performs close to Int-CPF.

The paper is organized as follows. Related work is reviewed in Section II, with system model and problem formulation presented in Section III. In Section IV, we present the *Coordinated Proportional Fairness (CPF)* allocation scheme for multi-association setting, and analyze its incentive compatibility. The integral variant of CPF (Int-CPF) and the selfish load balancing (SLB) scheme for single-association setting are presented and analyzed in Section V. In Section VI, we evaluate the performance of various schemes proposed. We conclude in Section VII.

II. RELATED WORK

A. Resource Allocation in Wireless Networks

The criterion of fairness has long played a central role in designing of resource allocation schemes. The most common understanding of fairness in computer networks is probably the *max-min fairness*, as defined in [7]: rates are made as equal as possible subject only to the constraints imposed by link capacities. However, *max-min fairness* is not an efficient resource allocation solution for elastic traffic in multi-rate wireless communication system, because when some MSs use a lower bit rate than the others, the performance of all MSs sharing the same BS is considerably degraded to the same level as the worst one, as shown in [8].

Compared to max-min fairness, proportional fairness as proposed by Kelly in [1] strikes a better balance between efficiency and fairness. Proportional fairness can be defined as the maximization of an objective function representing the overall utility of the flows in progress. The utility function chosen is logarithmic function of the allocated bandwidth, where the value of a flow for MS $m \in M$ increases with its allocated bandwidth R_m in proportional to $logR_m$. Formally, an allocation scheme S^* is proportional fair if and only if among all feasible schemes S:

$$S^* = \operatorname{argmax}_S \sum_{m \in M} \log R_m^{(S)} \tag{1}$$

Proportional fairness favors resource-efficient requests more than *max-min fairness*, by allowing large sharing to increase further with small sharing decreased, if change of the assigned bandwidth vectors result in the sum of the proportional changes to be non-negative. Thus, it helps improve system efficiency, while still preventing resource-efficient connections from starving resource-inefficient connections totally. In addition, it is shown by [9] to satisfy the axioms defining a Nash bargaining solution [10].

In a single-cell environment for both cellular networks [2] and Wi-Fi networks [3], the proportional fairness is implemented by allocating the radio resource of a BS (asymptotically) equally among associated MSs, regardless of their different efficiency in using the resource, i.e., their various link data rates. If timely channel feedback is available, channel-aware opportunistic scheduling algorithms [2] are often employed to exploit the "multi-user diversity". In this paper, we consider time-averaged channel state as input, and assume the underlying scheduling algorithm of each BS (which can be channel-aware) supports the resource allocation decision.

In a multi-cell wireless environment, techniques have been proposed to intelligently associate MSs with overlapping BSs to achieve globally optimal proportional fairness [4] [5] [6]. In this paper, we focus on the loadbalancing aspect of performance improvement, and assume the resource capacities of neighboring BSs are fixed and independent. The other forms of inter-cell optimization, such as dynamic channel assignment and interference avoidance, can be applied orthogonally.

B. Algorithmic Mechanism Design

Game theory aims to model situations in which multiple participants select strategies that have mutual consequences. Following the definitions used by Nisan et al. in [11], a game consists of a set of n players, 1, 2, ..., n. Each player i has his own set of possible strategies, say S_i . To play the game, each player i selects a strategy $s_i \in S_i$. We use $s = (s_1, ..., s_n)$ to denote the vector of strategies selected by the players and $S = \times_i S_i$ to denote the set of all possible ways in which players can pick strategies. The vector of strategies $s \in S$ selected by the players determines the outcome for each player. If by using a unique strategy, a user always gets better outcome than using other strategies, independent of the strategies played by the other players, we say that the strategy is the user's *dominant strategy*. If users select strategies such that, no player can unilaterally change its strategy to gain more payoff, we say that the game reaches a *Nash equilibrium*.

Algorithmic mechanism design [11] is a subarea of game theory which deals with the design of games. It studies optimization problems where the underlying data is *a priori unknown* to the algorithm designer, and must be implicitly or explicitly elicited from selfish participants (e.g., via a bid). The high-level goal is to design a protocol, or "mechanism", that interacts with participants so that *selfish behavior yields a desirable outcome*. Auction design is the most popular motivation in this area, though there are many others. When truth-telling is the dominant strategy of all participants, we say the mechanism is *incentive compatible*.



III. SYSTEM MODEL AND PROBLEM FORMULATION

Fig. 1: A convergent mobile communication system

Our discussion is based on a convergent system of heterogeneous wireless networks as shown in Figure 1. The main components of the considered architecture are: multi-mode terminals, all-IP core network, and the integrated radio access networks (RANs) sitting between them, as described below.

1) Ongoing silicon development enables chip makers to integrate multiple radio access technologies in a single chipset. For example, Qualcomm's Snapdragon chipset for mini-notebooks includes Wi-Fi alongside 3G, Bluetooth, broadcast TV and GPS (Global Positioning System) capabilities [12].

2) Meanwhile, wireless core networks are quickly evolving towards IP-based mechanisms [13]. IP layer enables provision of a richer set of services independent of the access networks.

3) As a bridge between the two components above, a flexible architecture capable of managing a large variety of coexisting radio access networks are being standardized [14] [15] [16]. The proposed Common Radio Resource Management (CRRM) functions consider the pool of resources in all radio access technologies (RATs) as a whole, aiming at better performance than stand-alone networks.

As shown in the figure, the radio resource manager can be interpreted as a logical entity which gathers input from different RATs, and coordinates resource allocation decisions among them. In practice, the channel state is measured by individual MS, which periodically feeds it back to the resource allocator for informed decision. Thus, an intelligent and selfish MS can manipulate its reported channel states, if it can gain more from network by doing so.

Based on this observation, a *multi-cell resource allocation procedure* can be interpreted as a game as follows. Consider a network with a set B of BSs and a set M of MSs. Let a link l = (m, b) be a pair of MS and BS that is able to communicate with each other. We call such a pair an *adjacent MS-BS pair*. Each MS $m \in M$ is a player of the game. The strategy of a MS m can be described as a channel state vector $R_m = (R_{mb}, b \in B)$, where R_{mb} gives the data rate supported between m and BS b. The resource allocation outcome is calculated according to the scheme employed by the resource manager, and the decision is enforced by individual BS. Note that, if the reported link data rate R_{mb} between MS m and BS b is not equal to the actual link data rate R_{mb}^* , the effective data rate will be less than R_{mb}^* . On one hand, if $R_{mb} < R_{mb}^*$, data is transferred by BS using R_{mb} . On the other hand, if $R_{mb} > R_{mb}^*$, data is transferred by BS at a rate higher than that can be fully decoded by MS, the resulted effective data rate becomes lower than that can be achieved by the most appropriate rate R_{mb}^* . As over-report can be easily detected [17], we focus on the case where m under-reports its channel state, i.e. $R_{mb} \leq R_{mb}^*$.

Formally, a multi-cell resource allocation game is defined as $(M, R^*, \mathbb{R}, S, x)$, where

- *M* is the set of MS players.
- $R^* = (R_m^*, m \in M)$ consists of the actual link data rate vector R_m^* for each MS $m \in M$.
- ℝ = ×_mℝ_m, m ∈ M, where ℝ_m = {R_m | R_m ≤ R^{*}_m} specifies the strategy space of MS m. m can choose
 any link data rate vector R_m ∈ ℝ_m when playing the game.
- S is an allocation scheme which determines the allocation vector based on the specified channel state input $R \in \mathbb{R}$.
- $x = (x_m, m \in M)$ gives the allocated data rate vector.

Applying the mechanism design framework [11], the *a priori* unknown underlying data in our game is the channel state experienced by individual MS. The algorithm designer (the resource allocator here) elicits the information through the periodic feedbacks of MSs. The high-level goal is to design a mechanism (the allocation scheme in our game), that interacts with participants so that selfish behavior yields a desirable outcome (an efficient and fair resource allocation in our game). Recall that, a mechanism is said to be *incentive compatible*, if the dominant strategy of each participant under the designed mechanism is to truthfully reveal its state. In our game, **incentive**

compatibility means the dominant strategy of each MS is to measure and report its channel state truthfully.

In contrast, if a game is not incentive compatible, MSs can gain by cheating about its state, thus making the system operate under inefficient state. Even worse, MSs may keep varying their behavior as response for others' strategies, which can lead to instability problem.

IV. MULTI-ASSOCIATION SETTING

This section presents a natural generalization of proportional fair allocation in an overlapping multi-cell environment, and analyzes its incentive compatibility. We assume each MS can simultaneously associate with multiple BSs to achieve aggregate throughput. We will consider the single-association setting in Section V.

A. Formulation

Given a link l, we use b(l) to denote the corresponding BS, and m(l) to denote the corresponding MS. We write L for the set of all links. If b = b(l), we set A_{bl} to be the required radio resource in BS b to support per unit flow through link l. If the channel condition between m(l) and b(l) is poor, it can only support a low data rate, thus more radio resource is required to transfer a unit of flow, which implies a higher resource consumption rate, i.e., A_{bl} is larger. On the other hand, if a MS-BS link is under good channel condition, less resource is required to transfer the same amount of data, thus, A_{bl} is smaller. As wireless channel state keeps changing with time, the value of A_{bl} used in our problem formulation is a time averaged link state which is relatively stable for a decision period. For $b \neq b(l)$, we set $A_{bl} = 0$, because sending flow over link l does not consume any resource of BS b. This defines a matrix $A = (A_{bl}, b \in B, l \in L)$.

For a given MS m, its several links through different BSs may substitute for one another. Formally, suppose that a MS m has a subset of L. We write $H_{ml} = 1$ if m = m(l), so that link l serves the MS m, and set $H_{ml} = 0$ otherwise. This defines a 0-1 matrix $H = (H_{ml}, m \in M, l \in L)$.

A flow pattern $y = (y_l, l \in L)$ supports the rates $x = (x_m, m \in M)$ if Hy = x, so that the flows over all links serving the MS m sum to the rate x_m . We let C_b be the finite radio resource capacity of BS b, for $b \in B$. A flow pattern y is feasible if $y \ge 0$ and $Ay \le C$, so that the resource consumed by wireless links through BS b sum to not more than its capacity. Note that we assume wireless transmissions are "orthogonal" (e.g., through time or code multiplexing), thus resource consumed by different links at the same BS can be linearly summed up, and resource usage in different BSs is independent of each other.

To illustrate the notations, we look at Figure 2. Each of MS m_1 and MS m_2 is equipped with both a cellular interface and a Wi-Fi interface. Both MSs locate in the overlapping coverage area of a Wi-Fi AP b_1 and a cellular BS b_2 . However, their channel conditions to the AP and BS are different. MS m_1 can communicate with Wi-Fi AP at 2Mbps and with cellular BS at 1Mbps, while MS m_2 can communicate with Wi-Fi AP at 1Mbps and with cellular BS at 2Mbps. There are 4 links corresponding to the 4 adjacent MS-BS pairs: $l_1 = (m_1, b_1)$, $l_2 = (m_1, b_2)$, $l_3 = (m_2, b_1)$, and $l_4 = (m_2, b_2)$. The input to CPF allocation problem is: MS set $M = \{m_1, m_2\}$, BS set



Fig. 2: CPF allocation example

$$B = \{b_1, b_2\}, \text{ link set } L = \{l_1, l_2, l_3, l_4\}, \text{ the matrix } A = \begin{bmatrix} \frac{1}{2} & 0 & 1 & 0\\ 0 & 1 & 0 & \frac{1}{2} \end{bmatrix}, \text{ and matrix } H = \begin{bmatrix} 1 & 1 & 0 & 0\\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

We assume unit capacity of both b_1 and b_2 , thus $C = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$.

Formally, the **Coordinated Proportional Fairness (CPF) allocation** is the optimal solution for the following problem:

$$\begin{array}{ll} maximize & \sum_{m \in M} w_m log(x_m) \\ s.t. & Hy = x, Ay \leq C \\ over & x, y \geq 0 \end{array} \tag{2}$$

where $w_m > 0$ is the weight assigned to different users representing their different priority. We consider only MSs with non-empty set of adjacent BSs, and BSs with non-empty set of adjacent MSs. Further, as $w_m log(0) = -\infty$ for all m, the optimal objective value for CPF allocation is achieved when Ay = C and x > 0. We can rewrite the constraints as follows without affecting the solution.

$$Hy = x, Ay = C, x > 0, y \ge 0$$
 (3)

The objective function is differentiable and strictly concave and the feasible region is compact. Thus, a maximizing value of (x, y) always exists and can be found by Lagrangian methods. There is a unique optimum for the rate vector x, since the objective function is a strictly concave function of x, but there may be many corresponding values of the flow rate y satisfying the constraints [18].

Let's look at the CPF allocation in the example of Figure 2. Its CPF solution is: $x = [2, 2]^T$, $y = [2, 0, 0, 2]^T$. The solution is Pareto-optimal. MS m_1 is served over link $l_1 = (m_1, b_1)$, and MS m_2 is served over link $l_4 = (m_2, b_2)$. Both m_1 and m_2 are assigned to the interface with more favorable channel.

By considering fairness in a global sense (among all MSs), the radio resource allocation solution automatically results in inter-cell load balance. Look at the example of Figure 2, if the channel condition between MS m_2 and BS b_2 deteriorates, and supports only a data rate of 0.8Mbps, BS b_2 becomes more congested than BS b_1 , in the sense that BS b_2 requires extra capacity in order to support the original allocation. The input for CPF problem becomes: $A = \begin{bmatrix} \frac{1}{2} & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{0.8} \end{bmatrix}$, with H unchanged. The CPF solution becomes: $x = [1.8, 0.9]^T$, $y = [1.8, 0, 0.1, 0.8]^T$, which automatically shifts some load of m_2 from b_2 to b_1 . Note that, the resource-efficient MS m_1 has a higher throughput than the resource-inefficient MS m_2 .

B. Incentive compatibility

Recall that a multi-cell resource allocation game is **incentive compatible** if the dominant strategy of each MS is to measure and report its channel state truthfully. Theorem 1 proves the positive result that in the *multi-cell resource allocation game* with CPF as the allocation scheme *S*, the dominant strategy for each MS is to report its channel state truthfully.

Theorem 1: A multi-cell resource allocation game with CPF allocation scheme is incentive compatible.

Proof: We prove this property by contradiction. Assume there is a user m^* which can increase its aggregate bandwidth allocation by not using truthful strategy. We denote the allocation decision for the original setting, where m^* does not cheat, as D' = (x', y'), and the allocation decision for the new setting, where m^* cheats, as D'' = (x'', y'').

Given a MS m, we denote the subset of its adjacent BSs that allocate strictly more radio resource to it in D'' than in D' as $B^+(m)$, i.e., $\forall b \in B^+(m), \frac{y'_{(mb)}}{R'_{mb}} > \frac{y'_{(mb)}}{R'_{mb}}$.

Given a BS b, we denote the subset of its adjacent MSs that get strictly lower radio resource allocation from it in D' than in D' as $M^{-}(b)$, i.e., $\forall m \in M^{-}(b), \frac{y'_{(mb)}}{R'_{mb}} < \frac{y'_{(mb)}}{R'_{mb}}$.

Denote the initial BS set as $B_0 = B^+(m^*)$. Based on our assumption, we have $x''_{m^*} > x'_{m^*}$. Thus, there must be some BSs which allocate more resource to m^* in D'' than in D'. More specifically, $B_0 \neq \emptyset$.

Denote the initial MS set as $M_0 = \bigcup_{b \in B_0} M^-(b)$. As a BS $b \in B_0$ allocates more resource to m^* in D'', and in both solutions D' and D'' it allocates all of its resources, it must reduce allocation to some other MS in D''. Thus, $M_0 \neq \emptyset$.

Consider the Lagrangian form of the CPF problem:

$$\mathbb{L}(x, y; \lambda, \mu)$$

$$= \sum_{m \in M} w_m log(x_m) - \lambda^T (x - Hy) + \mu^T (C - Ay)$$

$$= \sum_{m \in M} (w_m log(x_m) - \lambda_m x_m) + \sum_{l \in L} y_l (\lambda_{m(l)} - \mu_{b(l)} A_{b(l)l}) + \sum_{b \in B} \mu_b C_b$$
(4)

where $\lambda = (\lambda_m, m \in M), \mu = (\mu_b, b \in B)$ are vectors of Lagrange multipliers.

$$\frac{\partial \mathbb{L}}{\partial x_m} = (w_m log(x_m))' - \lambda_m \tag{5}$$

$$\frac{\partial \mathbb{L}}{\partial y_l} = \lambda_{m(l)} - \mu_{b(l)} A_{b(l)l}$$
(6)

Hence, at a maximum of \mathbb{L} , the following conditions hold:

$$\frac{w_m}{x_m} = \lambda_m$$

$$\lambda_{m(l)} = \mu_{b(l)} A_{b(l)l} \text{ if } y_l > 0$$
(7)

$$\leq \quad \mu_{b(l)} A_{b(l)l} \text{ if } y_l = 0 \tag{8}$$

The Lagrange multipliers λ and μ have simple interpretations. We may view μ_b as the implied cost of using unit radio resource of BS *b*, or alternatively the shadow price of adding additional radio resource at BS *b*. λ_m can be viewed as the weighted charge of unit flow for MS *m*.

As $x''_{m^*} > x'_{m^*}$, because of Equation 7, $\lambda''_{m^*} < \lambda'_{m^*}$. Thus, for any $b \in B_0$, because of Equation 8, $\mu''_b < \mu'_b$. Based on Equation 8 again, for any $m \in M_0$, $\lambda''_m < \lambda'_m$, thus $x''_m > x'_m$.

We repeatedly carry out the following set expansion step:

$$B_{n+1} = \bigcup_{m \in M_n} B^+(m) \cup B_n \tag{9}$$

$$M_{n+1} = \bigcup_{b \in B_{n+1}} M^{-}(b) \tag{10}$$

As B is a finite set, the process always terminates at some $n = n^*$ where $B_{n^*+1} = B_{n^*}$. For each expansion step, the argument about the change of Lagrange multipliers as in the initial step can still be applied, thus: $x''_m > x'_m, \forall m \in M_{n^*}$.

Consider B_{n^*} and M_{n^*} . For any MS $m \in M_{n^*}$, its allocated data rate strictly increases. For any MS $m \notin M_{n^*}$, its radio resource allocation from any BS $b \in B_{n^*}$ is not reduced according to the definitions above. Thus, BSs in B_{n^*} jointly allocate higher data rate in D'' to all MS $m \in M_{n^*}$ without affecting their allocation to any MS outside M_{n^*} . Combining the resource allocation decision of D'' for BSs in B_{n^*} and the allocation decision decision of D'for BSs not in B_{n^*} , we have a feasible allocation solution \tilde{x} for the original setting where m^* is honest. For MS m^* , \tilde{x}_{m^*} is the aggregate rate of m^* using actual link data rate, thus, we have $\tilde{x}_{m^*} \ge x'_{m^*} > x'_{m^*}$. For $m \in M_{n^*}$ and $m \neq m^*$, their reported data link rates are same for the two settings, thus $\tilde{x}_m = x''_m > x'_m$. Similarly, for $m \notin M_{n^*}$, $\tilde{x} \ge x'$. As the vector \tilde{x} is strictly larger than x', this contradicts with the fact that x' is Pareto optimal under the original setting where m^* is honest.

V. SINGLE-ASSOCIATION SETTING

The optimal solution for *CPF* allocation often requires MSs to be simultaneously assigned to multiple BSs, which may not be desirable in practice, due to the following reasons:

1) It requires a node to have multiple radios. On one hand, a software defined radio which can dynamically switch among different radio access technologies may not suffice, as it cannot simultaneously present in multiple overlapping cells. On the other hand, turning on multiple radios can significantly increase the power consumption.

2) When a single parameter changes in the network, the allocation decision may be adjusted globally. This may result in both system instability and excessive signaling overhead.

3) Transport protocol at client may have difficulty to efficiently aggregate the bandwidth from multiple interfaces, especially when the allocated bandwidth of each interface varies with time [19].

In this section, we study coordinated resource allocation in a *single-association setting*, where each MS is associated to a single BS.

A. Formulation and Complexity

The formulation for *CPF* allocation can be modified to reflect the additional constraint in single-association setting.

Formally, the **Integral Coordinated Proportional Fairness (Int-CPF) allocation** is the optimal solution for the following problem:

$$\begin{array}{ll} maximize & \sum_{m \in M} w_m log(x_m) \\ s.t. & Hy = x, Ay \leq C \\ & \forall m \in M, \exists l_m \in L, \forall l \neq l_m, H_{ml} y_l = 0 \\ over & x > 0, y \geq 0 \end{array}$$
(11)

Similar formulation has been proposed in [4] and [6] as well. If we decouple the solution for *Int-CPF allocation* scheme into the *inter-cell association control layer* and the *intra-cell scheduling layer*, we observe the optimal strategy for each BS in the second layer is independent of the association control decision in the first layer and the second layer strategy of each other. This is because one MS is served by a single BS, thus, every single BS should maximize the weighted logarithmic sum of data rate over MSs assigned to it, and it can achieve this by employing individual proportional fair scheduling. As the second level scheduling is clear, the remaining problem is to decide for each MS which BS it should associate to.

We show that, for *Int-CPF allocation* scheme, there does not exist an algorithm that can find the optimal solution in polynomial time unless P = NP, i.e., the problem is NP-hard. Similar to [20] and [4], our reduction is via *3-dimensional matching* problem which is known to be NP-complete. The *3-dimensional matching* problem is stated as follows.

Definition 1: Let $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_n\}$, $Z = \{z_1, \ldots, z_n\}$ be three disjoint sets with identical size n, and T is a subset of $X \times Y \times Z$. That is, T consists of triples (x, y, z) such that $x \in X$, $y \in Y$, and $z \in Z$. A $T' \subseteq T$ is a 3-dimensional matching if |T'| = n and $\bigcup_{t_i \in T'} t_i = B \cup C \cup D$. The problem is to find whether such a T' exists.

Theorem 2: Int-CPF allocation problem is NP-hard.

Proof: Consider a 3-dimensional matching problem where T consists of k triples (k > n), otherwise the problem becomes trivial). We construct a corresponding *Int-CPF allocation* problem as follows. For each tripe $t_i \in T$, we create a corresponding BS t_i with capacity 1. We create two types of MSs: normal MS and privileged MS. For each element $m \in X \cup Y \cup Z$, we create a corresponding normal MS m. There are totally 3n normal MSs. A normal MS *m* is covered by a BS t_i if and only if $m \in t_i$. In addition, we create k - n privileged MSs, which are covered by all BSs. We assume the link data rates of all adjacent MS-BS pairs are equal to a constant *R*. A normal MS has weight 1, while a privileged MS has weight $w_p > 2$. The weight of privileged MS is selected such that, if possible, packing the 3n normal MSs into *n* BSs, while assigning each of the k - n privileged MSs into each of the rest of k - n BSs, gives the highest value of $U_{max} = \sum_{m \in M} w_m log(x_m) = 3n log(\frac{R}{3}) + (k - n) w_p log(R)$. Thus, it is easy to verify that if there is a 3-dimensional matching solution *T'*, the *Int-CPF allocation* problem achieves the optimal solution U_{max} . Conversely, if the *Int-CPF allocation* problem achieves the optimal solution U_{max} , there is a 3-dimensional matching solution for the original problem.

Note that, if $w_m = 1, \forall m \in M$, and assuming that we know the congestion vector $(N_b, b \in B)$ for the optimal solution of Int-CPF, where N_b denotes the number of MSs assigned to BS b, we can reduce the problem of finding optimal solution for *Int-CPF allocation* to finding the maximum weight perfect k-matching in a bipartite graph as follows. Consider the complete bipartite graph G(M, B, E) where M denotes the MSs and B denotes the BSs. The requirements (k-values) of $m \in M$ is 1. For a BS $b \in B$, $k(b) = N_b$, the b^{th} coordinate of the congestion vector. The weight on each edge (m, b) is set to $w(m, b) = log(\frac{R_{mb}}{N_b})$, where R_{mb} is the link data rate between MS m and BS b. The optimal *Int-CPF allocation* corresponds to the maximum weight perfect k-matching, as each MS is associated with one BS, each BS gets the number of MSs as specified by the optimal congestion vector, and the logarithmic sum of allocated data rates for all MSs is maximized. Note that the number of possible congestion vectors is polynomial in |B|. In our evaluation, we use this approach to calculate the solution of *Int-CPF allocation* for a constant number of BSs.

B. Incentive Compatibility

In contrast to the multi-cell allocation game with *CPF allocation* scheme, the multi-cell allocation game with *Int-CPF allocation* scheme is not incentive compatible.

Theorem 3: A multi-cell resource allocation game with Int-CPF allocation scheme is not incentive compatible.

Proof: The theorem can be simply proved by providing counter examples.

In the example of Figure 3 (a), there is a MS m_1 covered by both the Wi-Fi AP b_1 and cellular BS b_2 , with rate of 1Mbps and $2 + \epsilon Mbps$ ($\epsilon > 0$) respectively. In addition, there is a MS m_2 which is covered only by b_2 with rate 1Mbps. The *Int-CPF allocation* is to assign MS m_1 to b_1 , and MS m_2 to b_2 , thus both m_1 and m_2 get a throughput of 1Mbps each. However, if MS m_1 cheats by hiding its association with b_1 , the *Int-CPF allocation* is to assign both MS m_1 and MS m_2 to b_2 , and allocate half the resource to each of them. In this case, m_1 gets a higher throughput of $1 + \frac{\epsilon}{2}Mbps$, while m_2 's throughput is reduced to $\frac{1}{2}Mbps$. This example shows that a MS with multiple adjacent BSs can manipulate its adjacent BS set, so as to be allocated to the BS that it prefers.

On the other hand, the example of Figure 3 (b) shows that, a MS can also manipulate its reported data rate to increase its benefit by changing other MS's association. In the given setting, both MS m_1 and MS m_2 are covered by both the Wi-Fi AP b_1 and cellular BS b_2 . Their link data rates to the AP and BS are shown in the figure. It is easy to verify that, the *Int-CPF allocation* is to assign MS m_1 to b_1 and m_2 to b_2 , such that the throughput of



(b) MS m₂ may cheat to gain

Fig. 3: Cheating under Int-CPF allocation

 m_1 is 1.5Mbps and the throughput of m_2 is 1Mbps. If MS m_2 cheats by hiding its adjacency with BS b_2 , i.e., set the data rate of link (m_2, b_2) to 0, the *Int-CPF allocation* will swap the assignment, with m_2 associated to b_1 while m_1 to b_2 . Thus, the throughput of m_2 increases to 1.4Mbps. Note that m_1 can react similarly by hiding its adjacency, which leads to vicious competition and results in both inefficiency and instability.

The example of Figure 3 (b) also shows that optimal *integral coordinated max-min fairness*, as proposed by Bejerano et. al. in [21], is also not incentive compatible.

C. Selfish Load Balancing: A Congestion Game Formulation

As *Int-CPF allocation* decision is computationally expensive, and does not define an incentive-compatible game, a natural alternative is to let selfish users decide for themselves which BS to associate with, while the coordinator just ensures that each MS is associated with a single BS at any time.

When each MS can make individual association decision directly, instead of the *multi-cell resource allocation* game as defined above, we have a *single-association game*.

A single-association game is defined as (M, S, x), where:

- *M* is the set of MS players.
- S = ×_mS_m denotes the set of all possible ways in which players can pick strategies. For each player m ∈ M, S_m denotes his own set of possible strategies, which corresponds to the subset of BSs which it can associate with. In particular, one strategy b_m ∈ S_m corresponds to the association of MS m with BS b_m. A strategy profile s ∈ S consists of the vector of each player's selected strategy, i.e. s = (b_m, m ∈ M).

Under the assumption that each BS implements individual proportional fairness, and that all users have the same weight, the throughput x_m received by a MS m under a strategy profile s = (b_m, m ∈ M) can be simply expressed as:

$$x_m(s) = \frac{R_{mb_m}}{N_{b_m}(s)} \tag{12}$$

where R_{mb_m} is the link data rate between MS m and its selected BS b_m , and $N_{b_m}(s)$ is the congestion level (number of associated users including MS m) of BS b_m under strategy profile s.

For each player in the *single-association game*, its reward (in terms of allocated bandwidth) of employing a certain strategy is affected only by the number of other players who employ the same strategy (choosing the same BS to associate with), rather than who they are. Thus, this game falls into the class of *congestion games* which is first introduced by Rosenthal in [22].

Rosenthal showed that if the reward function is the same for all players choosing the same strategy, then these games possess a rich structure, in particular they always have a Nash equilibrium in pure strategies. The term of "pure strategy" means each player *deterministically* plays a single chosen strategy, instead of randomly picks among multiple strategies. This result follows from the existence of a *potential function*, which is a real-valued function defined over the set of strategy profiles having the property that the gain of a player shifting to a new strategy is equal to the corresponding change of the potential function.

The existence of an exact potential function implies the *finite improvement property (FIP)*: Any sequence of strategy-tuples in which each strategy-tuple differs from the preceding one in only one coordinate (such a sequence is called a *path*), and the unique deviator in each step strictly increases the payoff he receives (an *improvement path*), is finite. The first strategy-tuple of a path is called the *initial point*; the last one is called the *terminal point*. Obviously, any *maximal improvement path*, an improvement path that cannot be extended, is terminated by a Nash equilibrium.

Milchtaich [23] extended the definition of *congestion game* to allow player-specific reward functions, i.e. different players have different rewards by choosing the same strategy. He showed that even these games have a pure Nash equilibrium.

In our setting, different MSs have different wireless link data rate with the same BS, thus the reward function is player-specific. However, the simple structure of the player-specific reward function as defined in Equation 12 allows us to prove a stronger result than Milchtaich. More specifically, Theorem 4 shows that the *single-association* game possesses *FIP*. To prove Theorem 4, we define for every strategy profile $s = (b_m, m \in M)$ the following potential function:

$$\Phi(s) = \sum_{m \in M} \log(R_{mb_m}) - \sum_{b \in B} \log(N_b(s)!)$$
(13)

where R_{mb_m} is the link data rate between MS m and its selected BS b_m , and $N_b(s)$ gives the number of MSs allocated to a BS b under the strategy profile s.

Theorem 4: Single-association game possesses the finite improvement property.

Proof: Consider a selfish step $s \to s'$ where a player $m \in M$ switches from BS b to BS b'. $\Phi(s') - \Phi(s) = (log(R_{mb'}) - log(N'_b(s) + 1)) - (log(R_{mb}) - log(N_b(s))) = log(x_m(s')) - log(x_m(s)).$

Based on the result, the *selfish load balancing (SLB)* scheme as described below always converges to a Nash equilibrium within finite steps. Under the *SLB* scheme, the resource manager starts from a feasible allocation decision, and a user is allowed to switch to a BS that can improve its throughput. In each iteration, only one user can switch. The iteration ends until a *Nash equilibrium* is reached, i.e., no user can unilaterally change its association to achieve a higher throughput.

Note that, there can be multiple Nash equilibria in the *single-association game*, and *SLB* can converge to any of them. For example, in Figure 3 (b), there are two Nash equilibria. In the first equilibrium, MS m_1 is associated to BS b_1 , while MS m_2 is associated to BS b_2 . In the second equilibrium, the associations are swapped. Individual MS can have significantly different bandwidth allocation under different Nash equilibria. Note that while *SLB* converges, it is not incentive compatible. For example, MS m_2 in Figure 3 (b) can hide its association with BS b_2 , so as to make the system converge to the second equilibrium instead of the first one.

Despite the fact that it is not incentive compatible, *SLB* is still a valuable solution, as no MS can gain by unilaterally change its association. In addition, our evaluation in Section VI shows that *SLB* converges quickly and performs close to *Int-CPF*. It remains an interesting research problem to design incentive-compatible resource allocation schemes for single-association setting, such that MS cannot gain by cheating, while system can operate in a fair and efficient state.

VI. EVALUATION

We compare the performance of the following six schemes. For allocation schemes which can split a user's flow among multiple interfaces, we consider:

- Coordinated Proportional Fairness (CPF) scheme, as formulated in Section IV-A.
- Uncoordinated Proportional Fairness (UPF) scheme, where a MS associates to all neighboring BSs by simultaneously turning on multiple radio interfaces. This represents a non-cooperative scenario where all users are selfish and the system is uncoordinated.

For single-association setting, which enforces each MS to associate with only a single BS, we consider:

- *Int-CPF* scheme as formulated in Section V-A. Despite its NP-hardness, we can use a maximum weight perfect k-matching formulation to calculate its decision in polynomial time when the number of BSs is a small constant as described in Section V-A.
- *Strongest-Signal-First (SSF)* scheme, which always associates a MS to the BS with the strongest received signal strength, regardless of network load. *SSF* scheme is the default association method for multiple radio access technologies, including Wi-Fi networks.
- Least-Population-First (LPF) scheme, which always associates a MS to the BS with the least number of associated MSs, regardless of the channel condition. LPF scheme is widely adopted in single-rate cellular networks.

• *Selfish Load Balancing (SLB)* scheme, as described in Section V-C. We use the allocation decision of *SSF* scheme as the initial point for *SLB* process.

For fair comparison, we assume each BS implements individual proportional fair scheduling for all schemes above, except for the case of *CPF* scheme, which decides for each BS its allocation vector, thus does not necessarily follow the individual proportional fair scheduling. We assume all MSs honestly report their channel state and association information.

Our simulation is based on a $600m \times 600m$ torus topology where 9 BSs are placed on a 3 by 3 grid, with the distance between two adjacent BSs set to 200 meters. All BSs have identical transmission power and operate on non-interfering channels. The maximum transmission range of a BS is set to 150 meters. The set B(m) of BSs covering a MS m are determined from MS m's location by examining whether its distance to a BS is within 150 meters. We have conducted evaluations for two user distributions:

- Uniform setting: MSs are distributed uniformly at random;
- *Hot spot setting*: Out of all MSs, 50% are randomly generated in a circle-shape hot spot with the radius of 150 meters around the center of a selected *hot BS*.

The arrival of MSs follows a Poisson process, and the sojourn time of a MS in the system follows an exponential distribution, both of which are independent of MSs' allocated throughput. Load $\rho = \frac{E[|M]|}{|B|}$ is defined as the average number of active MSs in the system divided by the number of BSs, with default value set to 5. We use the log-normal shadowing propagation model to calculate the received signal strength at MS from its adjacent BSs. Given the distance d < 150m between a MS m and a BS b, the received signal power $P_{dB}(d)$ from b at m is calculated as $P_{dB}(d) = P_{dB}(d_0) - 10\beta log_{10}\frac{d}{d_0} + X_{\sigma}$, where $d_0 = 10m$ is the reference distance, $\beta = 3$ is the path loss exponent, and X_{σ} is a Normal random variable in dB having a standard deviation of $\sigma = 12dB$ and zero mean. We set the Signal Noise Ratio (dB) within reference distance to $P_{dB}(d_0) - P_{dB}(N_0) = 35dB$, and use an empirical threshold-based mapping to determine the link data rate accordingly. The parameters are set to model the typical loss in an urban environment [24]. Each experiment below is carried out 500 times with different seeds, and the average is presented.

A. Performance comparison under honest report

Figure 4 (a) (b) plot the per-user throughput sorted in non-decreasing order under the uniform setting, and Figure 4 (c) (d) plot the result under the hot spot setting. Table I summarizes the arithmetic and geometric mean of per-user throughput under different schemes for the two settings respectively.

As expected, *Coordinated Proportional Fairness (CPF)* scheme produces the optimal geometric mean of peruser throughput. In contrary, *Uncoordinated Proportional Fairness (UPF)* performs much worse, providing arithmetic/geometric mean not only lower than *CPF* scheme, but also inferior to all other schemes except for *LPF* scheme in some cases. This observation holds for both uniform and hot spot settings. The significant performance gap between them strongly advocates the adoption of a coordinated resource allocation approach for integrated heterogeneous wireless networks.



	Uniform		Hot spot	
	Arith. Mean	Geo. Mean	Arith. Mean	Geo. Mean
CPF	1.26	0.89	1.21	0.81
UPF	0.89	0.67	0.87	0.59
Int-CPF	1.26	0.89	1.21	0.81
SLB	1.26	0.89	1.22	0.80
SSF	1.29	0.86	1.27	0.73
LPF	1.02	0.63	0.93	0.55

Fig. 4: Per-user throughput sorted in non-decreasing order

TABLE I: Arithmetic and geometric mean of per-user throughput (Mbps)

Among all allocation schemes for single-association setting, *Int-CPF* scheme has the optimal geometric mean of per-user throughput, given that users do not cheat. Its performance is close to CPF scheme under both uniform setting and hot spot setting. For both settings, the coordinate wise performance gap between the two schemes is never greater than 3.5%, and is less than 1% for more than 70% of MSs.

Our result also shows that, *Selfish Load Balancing (SLB)* scheme often has very close performance to *Int-CPF*, thus to *CPF* as well. For around 65% of user distribution in uniform setting, *SLB* scheme and *Int-CPF* make the identical association decision. For 99% of user distributions, the performance gap between the two schemes is less

than 1%. Similar phenomenon is observed under hot spot setting as well. In fact, such an approximation among *SLB*, *Int-CPF* and *CPF* holds for a wide range of settings, as demonstrated in Figure 5 which will be explained in detail later.

Based on this, we make the following observation: By using an (appropriate) single radio per user, the system can largely achieve the optimal performance when simultaneously using multiple radios per user.

Among all six schemes, *Strongest-Signal-First(SSF)* scheme achieves the highest arithmetic mean of per-user throughput. This is as expected, because *SSF* scheme greedily assigns each MS to the BS providing the best channel condition. However, *SSF*'s geometric mean of throughput is lower than *CPF*, *Int-CPF*, and *SLB* schemes, because a MS often associates to an overloaded BS which can only allocate a small portion of its overall radio resource to serve the MS, thus provide poor throughput despite of the high link data rate between them. This situation is particularly common in hot spot setting. As shown in Figure 4 (d), for hot spot setting, there are 15% of the MSs under *SSF* scheme have throughput lower than 150 Kilobit per second (Kbps), compared to 5% under CPF scheme. For the 15% of MSs with lowest bandwidth allocation, their throughput under *SSF* scheme is less than 70% of their throughput under *CCF* scheme. Thus, *SSF* can be unfair to a significant portion of users.

Least-Population-First(LPF) scheme often performs worst in terms of both arithmetic and geometric mean of per-user throughput, implying that traditional load balancing technique is not applicable to multi-rate wireless data networks.

Since geometric mean of per-flow throughput embodies both the overall system resource efficiency and fairness among users, in the following experiments we use geometric mean as the performance metric.



Fig. 5: Geometric mean of throughput under varying settings

Figure 5 (a) demonstrates different schemes' performance under varying load in uniform setting. The performance of *Int-CPF* and *SLB* scheme is close to each other in all range of load. Further, the performance gap between them and the *CPF* scheme reduces with increased traffic intensity. This is because both *Int-CPF* and *SLB* schemes allocate resource on a per MS basis. Hence, the larger the traffic load, the finer the relative granularity of them. In fact, the only case that we observe obvious difference between *CPF* scheme and the two single-association schemes is

when the average number of MS per BS is very small (e.g. \leq 3).

Figure 5 (b) demonstrates the impact of asymmetric traffic distribution. The figure shows that the performance of both *Int-CPF* scheme and *SLB* scheme is close to *CPF* scheme even under highly asymmetric traffic distribution. Such robustness is largely because both schemes take into account factors including both network load and link data rate. In comparison, performance of *SSF* deteriorates faster than all other schemes with increasing traffic asymmetry, while *LPF* scheme performs better than *UPF* scheme under high traffic asymmetry.

B. Strategic interaction under SLB and Int-CPF



Fig. 6: Convergence speed of SLB over varying load

Figure 6 shows the average number of steps required in the whole multi-cell system for *SLB* to converge to a Nash equilibrium, starting from the *SSF* allocation. As can be seen from the figure, *SLB* converges quickly, and the number of steps required grows linearly with the system load.

While *SLB* takes strategic interactions among MSs into consideration, *Int-CPF* simply ignores them. Our evaluation also shows that, up to 15% of decision made by *Int-CPF* is not a Nash equilibrium in the single-association game. More specifically, there is at least one user who can unilaterally changes its association to gain higher throughput from the network under *Int-CPF* allocation. As illustrated in Figure 3 (a), the user can cheat by hiding all of its adjacent BSs except for its desired BS, so as to affect the *Int-CPF* resource allocation decision and increase its own bandwidth.

We observe that, in more than 30% of settings, there are multiple Nash equilibria in the induced *single-association game*, meaning that there is MS that can cheat about its channel state to drive the system to the Nash equilibrium that it prefers.

VII. CONCLUSION

This paper considers the *coordinated radio resource allocation problem* for users which are simultaneously covered by multiple overlapping heterogeneous wireless networks. We introduce incentive compatibility as a new criterion for making allocation decision by formulating the resource allocation process as a *multi-cell resource allocation game*. We prove that the game with CPF allocation is incentive compatible. However, for the single-association setting, the integral version of CPF allocation (Int-CPF) is both computationally expensive and prone

to user-manipulation. Alternatively, we propose the selfish load balancing scheme, which can quickly converge to a Nash equilibrium, and often achieves performance near to CPF allocation as shown by our simulation.

REFERENCES

- [1] F. Kelly, "Charging and rate control for elastic traffic," European Transactions on Telecommunications, vol. 8, no. 3, pp. 33–37, 1997.
- [2] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users," *IEEE Communications Magazine*, vol. 38, no. 7, pp. 70–77, July 2000.
- [3] G. Tan and J. Guttag, "Time-based fairness improves performance in multi-rate WLANs," in USENIX, Boston, MA, USA, June 2004.
- [4] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *IEEE INFOCOM*, Barcelona, Spain, April 2006.
- [5] B. Chen and M. Chan, "Proportional fairness for overlapping cells in wireless networks," in VTC 2006-Fall, Montreal, Canada, September 2006.
- [6] L. Li, M. Pal, and Y. Yang, "Proportional fairness in multi-rate wireless LANs," in INFOCOM, Phoenix, AZ, USA, April 2008.
- [7] D. Bertsekas and R. Gallager, Data Networks (Second edition). Prentice Hall, 1992.
- [8] M. Heusse, F. Rousseau, G. Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *IEEE INFOCOM*, San Francisco, CA, USA, March 2003.
- [9] R. Mazumdar, H. Yaiche, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Trans. on Networking*, vol. 8, no. 5, pp. 667–677, 2000.
- [10] J. Nash, "The bargaining problem," Econometrica, vol. 18, no. 2, pp. 155-162, April 1950.
- [11] N. Nisan, T. Roughgarden, E. Tardos, and V. V. (editors)), Algorithmic game theory. Cambridge University Press, 2007.
- [12] Qualcomm, "Snapdragon chipset," http://www.qctconnect.com.
- [13] S. Sesia, I. Toufik, and M. B. (editors), LTE The UMTS Long Term Evolution From Theory to Practice. John Wiley and Sons, 2009.
- [14] 3GPP, "Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking," 3GPP TR 22.934 V6.2.0, September 2003.
- [15] —, "Requirements on 3GPP system to Wireless Local Area Network (WLAN) interworking (Release 6)," 3GPP TS 22.234 V6.0.0, March 2004.
- [16] EVEREST, "Deliverable 11: First report on the evaluation of RRM/CRRM algorithms," November 2004.
- [17] S. Zhong, L. Li, Y. Liu, and Y. Yang, "On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks — an integrated approach using game theoretical and cryptographic techniques," in *MobiCom*, Cologne, Germany, August 2005.
- [18] P. Whittle, Optimization Under Constraints. Wiley, Chichester, 1971.
- [19] H. Hsieh and R. Sivakumar, "A transport layer approach for achieving aggregate bandwidths on multi-homed mobile hosts," in *MobiCom*, Atlanta, GA, USA, 2002.
- [20] J. Lenstra, D. Shmoys, and E. Tardos, "Approximation algorithms for scheduling unrelated parallel machines," *Mathematical Programming*, vol. 46, no. 1, pp. 259–271, January 1990.
- [21] Y. Bejerano, S. Han, and L. Li, "Fairness and load balancing in wireless LANs using association control," in *MobiCom*, Philadelphia, PA, USA, September 2004.
- [22] R. Rosenthal, "A class of games possessing pure-strategy nash equilibria," International Journal of Game Theory, vol. 2, pp. 65–67, 1973.
- [23] I. Milchtaich, "Congestion games with player-specific payoff functions," *Games and Economic Behavior*, vol. 13, no. 1, pp. 111–124, 1996.
- [24] T. Rappaport, Wireless Communications: Principles and Practice (2nd Edition). Prentice Hall PTR, 2001.