Optimal intercell coordination for multiple user classes with elastic traffic

Prajwal Osti, Pasi Lassila, Samuli Aalto
Aalto University School of Electrical Engineering,
Department of Communications and Networking,
Otakaari 5, FI-02150, Finland
Email: firstname.lastname@aalto.fi

***Abstract*—We consider the intercell coordination problem between two neighboring cells, assuming that the traffic in the system consists of elastic downlink data flows. In this case, there is an option of completely switching off one base station at certain times, which reduces interference and enables a higher service rate in the neighboring base station. We use a flow level queueing model to describe the evolution of the system based on a symmetric capacity region. Recent results by Verloop and Núñez-Queija show that, assuming a single class of flows for each cell, the stochastically optimal dynamic policy is to have both stations switched on whenever there are users in both cells. In this paper, we consider a system where the two stations are able to provide services to two different classes of users — the near ones and the far ones. In this setting, the stochastic optimality of the both stations on policy does not necessarily hold, but it may still be a close-to-optimal policy, at least for minimizing the mean flow delay. We present a systematic method based on the policy improvement algorithm of the theory of the Markov Decision Processes to generate a near-optimal state-dependent resource allocation policy. Our numerical experiments with these two dynamic policies indicate that the both stations on policy is, indeed, close to optimal even when there are multiple user classes.**

***Index Terms*—cellular networks, intercell coordination, elastic traffic, flow-level models**

I. Introduction

The capabilities of modern cellular systems, such as HSPA and 4G, enable the development of sophisticated mechanisms for mitigating the impacts of interference from simultaneous transmissions coming from neighboring cells. Intercell coordination refers to the idea that different base stations, instead of operating independently, work together in a coordinated fashion to enhance the performance of the whole system based on some global information of the system. For example, a base station may take into account the traffic in neighboring cell, the power being transmitted by other stations, the size and the type of the flows being served, *etc*., before deciding on its own service strategy. However, determining optimal resource allocation strategies in a multicell setting also poses significant challenges.

We assume that the traffic in the system consists of elastic downlink data flows. In this case, one has even more flexibility in coordination activities as there is an option of completely switching off some base stations at certain times. This reduces interference and enables a higher service rate in the neighboring base stations. At the flow level, the impact of such resource allocation policies on the service rates of the system can be formalized by the notion of *the capacity region*. The capacity region characterizes the set of all possible long-term average service rates that can be realized by the resource allocation in order to serve the flows.

The capacity region abstraction of the system allows the development of queuing theoretic models to analyze the structural properties of resource allocation policies. The base stations correspond to servers and if the base stations are independent, the system can be modeled using classical M/G/1 queueing models, see [1], [2]. However, as in our case, turning the stations on/off depending on the state of the system makes the service rate of the stations coupled, and it is reflected in the corresponding capacity region of the system. In general, for such models even the fundamental notion of stability is notoriously difficult to characterize. Indeed, most of the work focuses on these stability properties, see, *e.g.*, [3]–[5]. Results on optimizing the resource allocation to minimize the flow delay are significantly more scarce.

Recently, structural results on the optimal resource allocation in a specific case with two base stations have been given by Verloop and Núñez-Queija in [6]. They consider a system with *asymmetric* service rates, where flows arrive at the stations according to a Poisson process and the service requirements are exponentially distributed. One of the results states that under a specific condition for the service requirements and the asymmetric rates, the stochastically optimal dynamic policy is to have both stations switched on whenever there are users in both cells. We also study the same system with two interfering base stations under the assumption of a *symmetric* capacity region and symmetric service requirements. More specifically, when only one station is on it serves at a rate c_0 (the same for both stations) and if they are both on, they both serve at a rate $c_1 < c_0$. Moreover, we focus on the so-called *low interference* case that is characterized by the assumption that $c_0 < 2c_1$. Under this assumption, the both stations on policy maximizes jointly the speed at which the system operates and the rate at which flows complete. Thus, the stochastic optimality of the both stations on policy is intuitively quite easy to understand.

In this paper, we extend the idea of [6] and consider a system where the two stations are able to provide services to two different classes of users — the near ones and the far ones. We model these user types by different classes of a multi-class processor sharing system so that the far off users are able to achieve only a fixed fraction of the service rate of the near users. In this setting, the stochastic optimality of the both stations on policy does not necessarily hold. For example, consider a situation where in one cell there are only near users and in the other one only far off users. Then, with a suitable choice of parameters, the both stations on policy no longer jointly maximizes the speed at which the system operates and the rate at which flows complete.

The both stations on policy may still be a close-to-optimal policy, at least for minimizing the mean flow delay, and our objective is to gain insight into if this is the case. To this end, the theory of Markov Decision Processes and the policy improvement algorithm is used to determine an improved dynamic policy after the first iteration step when starting from the optimal static policy. Typically, after the first iteration, the resulting policy is already close to optimal. Comparing then this improved dynamic policy to the both stations on policy yields an indication of how close the both stations on policy is to the optimal one. For the construction of the improved dynamic policy, we utilize the results given by Leino and Virtamo in [7]. It turns out that this policy is characterized by linear switching curves in the four-dimensional state space.

In our numerical experiments, we compare the performance of the both stations on policy and the improved dynamic policy to the optimal static policy by means of simulations. The results provide clear evidence of the near-optimality of the both stations on policy, at least in the mean value sense. When the load is symmetric, the both stations on policy is better than the optimal static policy by an almost constant factor as the load is uniformly increased, and also outperforms consistently the improved dynamic policy. In the extreme case with only near users in one cell and far off users in the other, we can force the system to a situation where the both stations on policy may suffer, as discussed earlier. However, even in this case the both stations on policy yields virtually identical performance to the improved dynamic policy at lower loads and at higher loads the both stations on policy actually becomes better.

The rest of the paper is organized as follows. In Section II, we describe in detail the model used. The optimal static policy is derived in Section III. This is then followed by Section IV, where the Markov Decision Process framework is given and the improved dynamic policy is constructed. Numerical results are given and discussed in Section V. Finally, conclusions are presented in Section VI.

II. System model and problem formulation

Consider a two-cell system in which each cell and the base station in it are indexed by the elements of $\mathcal{I} = \{1, 2\}$ and let $i \in \mathcal{I}$ unless otherwise stated. Each base station can operate at the maximum rate of c_0 (bits/sec) when the other base station is turned off. When both stations are switched on, each station is capable of operating at maximum rate of c_1 (bits/sec), where $c_1 < c_0$ due to interference between the two cells. In addition, we assume that $c_0 < 2c_1$, which is referred to as the low interference case.

By means of time sharing, any rate pair $\boldsymbol{r} = (r_1, r_2) \in \mathcal{C} = \textbf{Conv}\{(0,0),(c_0,0),(0,c_0),(c_1,c_1)\}$ can be allocated to the two base stations in the long run, where $\textbf{Conv}\, A$ is the convex hull of the elements of A. The capacity region $\mathcal{C}$ is clearly the intersection of the following half planes (as illustrated in Figure 1):

$$r_1 \geq 0 \tag{1}$$

$$r_2 \geq 0 \tag{2}$$

$$r_2 \leq c_0 - \frac{c_0 - c_1}{c_1} r_1 \tag{3}$$

$$r_2 \leq \frac{c_1}{c_0 - c_1}(c_0 - r_1). \tag{4}$$

In each cell, we have two classes of users, the *near users* and the *far off users*. Due to the path loss effects, the far off users can get only a fraction of the data rate that the near users get. More specifically, if a base station can serve the near users at the rate r, then it can give a data rate of only kr, $k < 1$, to the far off users. For each cell i, we assume that the near users arrive at rate λ_{i1} and the far off users at rate λ_{i2}. Traffic consists of elastic flows (such as TCP file transfers) downloaded by

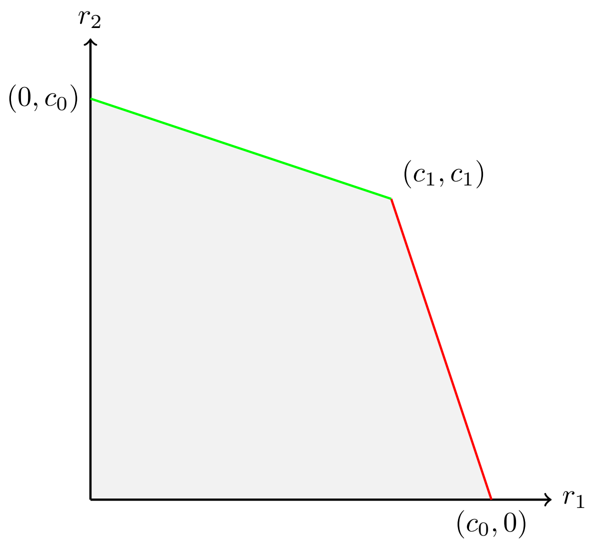

Fig. 1: Capacity region $\mathcal{C}$ when $c_0 < 2c_1$

the users (one flow per user). The arrivals are assumed to be from independent Poisson processes. In addition, we assume that the flow sizes are exponentially distributed with mean $\mathbf{E}[X] = \bar{X}$ (bits). Each base station shares its transmission time fairly among the users it is serving, working as a two-class *processor-sharing* (PS) system.

Let $\boldsymbol{x} = (x_1, x_2)$ denote the number of users present in Cell 1 and $\boldsymbol{y} = (y_1, y_2)$ denote the number of users present in Cell 2. In both cases, the first element of the vector represents the number of the near users while the second element represents the number of the far off users in the cell. We are interested in state-dependent *resource allocation policies* $\pi \in \Pi$ described by the allocation function $\pi(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{C}$. An example is the *Both Stations On* policy, which keeps both stations on whenever they are non-empty, defined by

$$\pi(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} (c_1, c_1), & \text{for } \boldsymbol{x} \neq (0,0) \text{ and } \boldsymbol{y} \neq (0,0); \\ (c_0, 0), & \text{for } \boldsymbol{y} = (0,0); \\ (0, c_0), & \text{for } \boldsymbol{x} = (0,0). \end{cases}$$

In [6], the Both Stations On policy is shown to be (i) maximally stable and (ii) even stochastically optimal (minimizing the total number of users in the two cells in stochastic sense) under the assumption that there is only one class per station. As speculated in Introduction, it is no longer clear how optimal the Both Stations On policy is in our case with two classes per cell. Below in Sections III and IV, we present a systematic method based on the policy improvement algorithm of the theory of the Markov Decision Processes that can be used to generate a near-optimal state-dependent resource allocation policy, which we will call the *Improved Dynamic* policy. The two policies are finally compared in Section V.

III. Optimal static policy

In this section we consider *static policies* under which the system is operated with a *fixed* rate pair $\boldsymbol{r} = (r_1, r_2) \in \mathcal{C}$ (independently of the state of the system). We derive the *Optimal Static* policy $\boldsymbol{r}^*$ that minimizes the total average number of users in the system (as well as the mean flow delay due to Little's formula).

Since the service rate allocations for the two stations are fixed with a static policy, each station can be modelled as an independent M/G/1-PS queue. The traffic intensity in station i is $\rho_i \triangleq \frac{\lambda_{i1}\bar{X}}{r_i} + \frac{\lambda_{i2}\bar{X}}{kr_i}$, and the average number of users in each station is given by

$$\mathbf{E}[N_i] = \frac{\rho_i}{1 - \rho_i}. \tag{5}$$

The average aggregate number of users in the whole system is then

$$\mathbf{E}[N] = \mathbf{E}[N_1] + \mathbf{E}[N_2] = \frac{\rho_1}{1-\rho_1} + \frac{\rho_2}{1-\rho_2} \tag{6}$$
$$= \frac{\alpha_1}{r_1 - \alpha_1} + \frac{\alpha_2}{r_2 - \alpha_2}, \tag{7}$$

where $\alpha_i \triangleq \bar{X}(\lambda_{i1} + \lambda_{i2}/k)$. These parameters α_i can be interpreted as the total arrival rates of *bits* at the base station.

For the most efficient use, the rate pair is always taken from the border of the capacity region, *i.e.*,

$$r_2 = \begin{cases} c_0 - \frac{c_0 - c_1}{c_1} r_1 & \text{if } 0 \leq r_1 \leq c_1, \\ \frac{c_1}{c_0 - c_1}(c_0 - r_1) & \text{if } c_1 \leq r_1 \leq c_0. \end{cases} \tag{8}$$

Moreover, the allocated service rates must always keep the system stable, *i.e.*, $r_i > \alpha_i$ for each $i \in \{1, 2\}$.

The optimal static operating policy can be found by minimizing (7) over all the possible points in the capacity region where the system is stable. Without loss of generality, we assume here that the aggregate arrival rate of bits in Cell 2 is greater than that in Cell 1, *i.e.*, $\alpha_2 \geq \alpha_1$.

If $\alpha_2 \geq c_1$, a portion of the upper boundary region only (and no part of the right boundary region) is a subset of stability region, *i.e.*, $0 \leq r_1 \leq c_1$ and

$$r_2 = c_0 - \frac{c_0 - c_1}{c_1} r_1. \tag{9}$$

For minimum, $\frac{\mathrm{d}\mathbf{E}[N]}{\mathrm{d}r_1} = 0$, from which it straightforwardly follows that

$$r_1^* = \frac{\frac{c_1}{c_0 - c_1}(c_0 - \alpha_2)\sqrt{\alpha_1} + \alpha_1\sqrt{\alpha_2 \frac{c_1}{c_0 - c_1}}}{\sqrt{\alpha_1} + \sqrt{\alpha_2 \frac{c_1}{c_0 - c_1}}}, \tag{10}$$

where r_1^* is the value of r_1 that minimizes (6). The corresponding r_2^* can be calculated from (9).

When $\alpha_2 < c_1$, the right boundary of the stability region also has stable rate pairs, and possibly an optimal rate pair. But as the average number of users is an increasing function of r_1 in this region, the optimal value should exist either at the common point of the left and the top boundaries, (c_1, c_1), or at the upper boundary which is given by (10).

So the optimal static rate pair for the stable arrival rates is determined by the simple procedure highlighted below:

1) The optimal r_1^* and the corresponding r_2^* are calculated from (10) and (9).
2) If $r_1^* \leq c_1$, then this is taken as the optimal rate. Otherwise (c_1, c_1) is the optimal rate.

These optimality results are obtained assuming that the service rate allocations for the two base station are static and that it cannot change as the system evolves. However, one can look for a *dynamic policy* that checks the system at certain instants of time and, depending on the state of the system at that time, assigns a new service rate pair. Such a policy can give better performance than the optimal static policy.

IV. Improved dynamic policy

In this section we apply the *policy improvement* algorithm of the theory of Markov Decision Processes, originally due to Howard [8]. To run the algorithm, we need an initial policy for which we are able to calculate so-called relative values of states. Below in Section IV-A, we apply the results derived by Leino and Virtamo [7] to determine the relative values of states for the optimal static policy. An improved dynamic policy can then be constructed by applying the policy improvement algorithm, as demonstrated in Section IV-B. A similar approach (called First Policy Iteration) has recently been used by Hyytiä *et al.* [9] to find near-optimal size-aware dispatching rules for parallel queueing systems.

A. Relative values of states for multi-class PS queues

Consider a queueing system with one server which is serving C classes. The jobs for class i arrive according to an independent Poisson process at rate λ_i and have exponentially distributed sizes with mean $1/\mu_i$. The traffic load of the i-th class, ρ_i, is defined as $\rho_i \triangleq \frac{\lambda_i}{\mu_i}$ and the traffic intensity of the whole system ρ defined as $\rho \triangleq \sum_{i=1}^{C} \rho_i$. We consider a stable system with $\rho < 1$. The server processes the jobs according to the PS discipline. These assumptions imply that the system can be fully described as a Markov process, $X(t)$, which takes the values in a C-dimensional state-space $\mathcal{X} \subseteq \mathbb{Z}_+^C$, where the state $\boldsymbol{x} = (x_1, \ldots, x_C) \in \mathcal{X}$ represents the number of customers, x_i, in class i.

Each job in the system generates holding costs with rate 1. Thus, when $X(t) = \boldsymbol{x}$, the *instantaneous cost rate* $r(\boldsymbol{x})$ is equal to $r(\boldsymbol{x}) = \sum_{i=1}^{C} x_i$. In addition, the average cost rate $\bar{r}$ is clearly equal to $\bar{r} = \rho/(1-\rho)$. The *relative value of state*, $v(\boldsymbol{x})$, describes the expected cumulative difference in costs between a system starting from state $\boldsymbol{x}$ and a stationary system with the initial distribution equal to the equilibrium distribution,

$$v(\boldsymbol{x}) = \mathbf{E}\left[\int_0^\infty (r(X(t)) - \bar{r})\ \mathrm{d}t \ \middle|\ X(0) = \boldsymbol{x}\right].$$

These relative values of states satisfy the so-called *Howard equations*:

$$r(\boldsymbol{x}) - \bar{r} + \sum_{\boldsymbol{y} \in \mathcal{X}} q_{\boldsymbol{x}\boldsymbol{y}}(v(\boldsymbol{y}) - v(\boldsymbol{x})) = 0 \quad \forall \boldsymbol{x} \in \mathcal{X},$$

where $q_{\boldsymbol{x}\boldsymbol{y}}$ is the transition rate from state $\boldsymbol{x}$ to state $\boldsymbol{y}$. For the multi-class PS queueing system, the Howard equation for state $\boldsymbol{x} = (0, \ldots, 0)$ reads as

$$-\bar{r} + \sum_{i=1}^{C} \lambda_i (v(\boldsymbol{x} + \boldsymbol{e}_i) - v(\boldsymbol{x})) = 0, \tag{11}$$

and for states $\boldsymbol{x} \neq (0, \ldots, 0)$ we have

$$\sum_{i=1}^{C} x_i - \bar{r} + \sum_{i=1}^{C} \lambda_i (v(\boldsymbol{x} + \boldsymbol{e}_i) - v(\boldsymbol{x})) + \sum_{i=1}^{C} \frac{\mu_i x_i}{\sum_{j=1}^{C} x_i} \big(v(\boldsymbol{x} - \boldsymbol{e}_i) - v(\boldsymbol{x})\big) = 0. \tag{12}$$

Here $\boldsymbol{e}_i$ denotes the unit vector into corresponding to class i.

As shown in [7], the relative values of the states $v(\boldsymbol{x})$ for the multi-class PS queue are given by the following quadratic expression:

$$v(\boldsymbol{x}) = \sum_{i=1}^{C} \sum_{j=1}^{C} a_{ij} x_i x_j + \sum_{i=1}^{C} a_i x_i + a_0, \tag{13}$$

where $a_{ij} = a_{ji}$. Then with proper substitutions followed by some algebraic manipulations, equations (12) become

$$\sum_{i=1}^{C}\left(1+2\sum_{j=1}^{C}\lambda_j a_{ji}-2\mu_i a_{ii}\right)x_i^2 + 2\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\left[1+\sum_{k=1}^{C}\lambda_k(a_{ki}+a_{kj}) -\mu_i a_{ij}-\mu_j a_{ji}\right]x_i x_j - \sum_{i=1}^{C}\left[\mu_i(a_i-a_{ii})\right]x_i = 0 \quad \text{for all } \boldsymbol{x}\neq(0,\ldots,0). \tag{14}$$

The left-hand side of (14) is a quadratic polynomial in C variables without the constant term with a total of $C+\frac{C(C-1)}{2}+C=\frac{C^2+3C}{2}$ terms. These equations are satisfied by any state of the underlying Markov process implying that the coefficients of all the terms of the polynomial are zero. So we get $\frac{C^2+3C}{2}$ linear equations in a_i and a_{ij}, $i,j\in\{1,\ldots,C\}$ with $a_{ij}=a_{ji}$, *i.e.*, again $\frac{C^2+3C}{2}$ unknowns. We can easily see, from the coefficients of the linear terms of x_i, that $a_i=a_{ii}$. Considering the fact that the coefficients of the quadratic terms of (14) do not depend on a_i, only $\frac{C^2+C}{2}$ linear equations corresponding to the coefficients of the quadratic terms — x_i^2 and $x_i x_j$ — can be solved to get all a_i and a_{ij} where $i,j\in\{1,\ldots,C\}$.

For the case of two classes ($C=2$) we get the following quadratic equation:

$$\begin{aligned}&(1+2\lambda_1 a_{11}+2\lambda_2 a_{21}-2\mu_1 a_{11})x_1^2\\&+(1+2\lambda_1 a_{12}+2\lambda_2 a_{22}-2\mu_2 a_{22})x_2^2\\&+2[1+\lambda_1(a_{11}+a_{12})+\lambda_2(a_{21}+a_{22})\\&-\mu_1 a_{12}-\mu_2 a_{21}]x_1x_2\\&-\left[\mu_1(a_1-a_{11})\right]x_1-\left[\mu_2(a_2-a_{22})\right]x_2=0.\end{aligned}$$

Thus, there are five independent linear equation in as many unknowns:

$$\begin{aligned}&1+2\lambda_1 a_{11}+2\lambda_2 a_{21}-2\mu_1 a_{11}=0,\\&1+2\lambda_1 a_{12}+2\lambda_2 a_{22}-2\mu_2 a_{22}=0,\\&1+\lambda_1(a_{11}+a_{12})+\lambda_2(a_{21}+a_{22})\\&\quad-\mu_1 a_{12}-\mu_2 a_{21}=0,\\&\mu_1(a_1-a_{11})=0,\\&\mu_2(a_2-a_{22})=0,\end{aligned}\tag{15}$$

resulting in

$$\begin{aligned}a_{11}&=a_1=\frac{1}{G}\left(\frac{\mu_2}{\mu_1}+1-\rho\right),\\a_{22}&=a_2=\frac{1}{G}\left(\frac{\mu_1}{\mu_2}+1-\rho\right),\\a_{12}&=\frac{1}{G}(2-\rho),\end{aligned}\tag{16}$$

where $G=2(1-\rho)(\mu_1+\mu_2-\lambda_1-\lambda_2)$.

Note that the parameter a_0 cannot be determined by this way. On the other hand, it does not matter since a_0 is finally not needed for the policy improvement algorithm, as we will see below.

B. Policy improvement

Once the relative values of states are determined for an initial policy, a policy improvement step can be performed to determine a better policy. Below we first describe the method in a general queueing setting, and then apply the method to our specific queueing system.

Let $\mathcal{X}$ refer to the set of all states $\boldsymbol{x}$. We define $A(\boldsymbol{x})$ as the set of all the actions that can be chosen when the system is in state $\boldsymbol{x}$. For the policy improvement algorithm, we need an *initial policy* $\vec{\pi}=(\pi_{\boldsymbol{x}})$ with state-dependent actions $\pi_{\boldsymbol{x}}\in A(\boldsymbol{x})$. Then we construct an improved policy $\vec{\pi}'=(\pi'_{\boldsymbol{x}})$ by choosing the action $\pi'_{\boldsymbol{x}}$ in state $\boldsymbol{x}$ such that

$$\pi'_{\boldsymbol{x}}=\underset{a\in A(\boldsymbol{x})}{\arg\min}\left\{r_{\boldsymbol{x}}(a)-\bar{r}(\vec{\pi})+\sum_{\boldsymbol{y}\in\mathcal{X}}q_{\boldsymbol{xy}}(a)v_{\boldsymbol{y}}(\vec{\pi})\right\},\tag{17}$$

where $r_{\boldsymbol{x}}(a)$ is the instantaneous cost rate when action a is chosen, $\bar{r}(\vec{\pi})$ is the average cost rate when the initial policy $\vec{\pi}$ is used, $q_{\boldsymbol{xy}}(a)$ is the transition rate from state $\boldsymbol{x}$ to state $\boldsymbol{y}$ when action a is chosen, and $v_{\boldsymbol{y}}(\vec{\pi})$ is the relative value of state $\boldsymbol{y}$ under the initial policy.

We choose our initial policy as the optimal static policy and then apply the policy improvement algorithm once. As the system under the initial policy consists of two independent two-class PS queues, we can construct the relative values of states for each queue from the discussion in Section IV-A. These relative values can then be added together to get the relative values for the whole system. The policy improvement process consists of choosing one of three possible actions, *viz.*, 10, 01 and 11, which correspond to only Station 1 being on, only Station 2 being on, and both stations being on respectively, that leads to the smallest expected costs in the future.

As before, let $\boldsymbol{x}=(x_1,x_2)$ denote the number of users present in Cell 1 and $\boldsymbol{y}=(y_1,y_2)$ denote the number of users present in Cell 2. Since any static policy results in two independent multiclass PS queues, we conclude that the relative value of state, $v_s(\boldsymbol{x},\boldsymbol{y})$, when the optimal static policy is used is given by,

$$v_s(\boldsymbol{x},\boldsymbol{y}) = v_{1s}(\boldsymbol{x}) + v_{2s}(\boldsymbol{y}),$$

where $v_{1s}(\cdot)$ and $v_{2s}(\cdot)$ are the relative values of states for Cells 1 and 2, respectively, for the corresponding optimal static policies. It follows from Section IV-A that

$$v_{1s}(\boldsymbol{x}) = a_{11}x_1(x_1+1) + a_{22}x_2(x_2+1) + 2a_{12}x_1x_2 + a_0,$$
$$v_{2s}(\boldsymbol{y}) = b_{11}y_1(y_1+1) + b_{22}y_2(y_2+1) + 2b_{12}y_1y_2 + b_0.$$

The coefficients a_{ij} and b_{ij} are the appropriate relative value parameters for Cells 1 and 2, respectively, which are determined by (16).

With the optimal static policy used as the initial policy, the expressions in (17) for the three possible actions $a \in \{10, 01, 11\}$ are as follows:

$$\begin{aligned} r_s(\boldsymbol{x},\boldsymbol{y}) - \bar{r}_s &+ \lambda_{11}[v_s(\boldsymbol{x}+\boldsymbol{e}_1,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ &+ \lambda_{12}[v_s(\boldsymbol{x}+\boldsymbol{e}_2,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ &+ \lambda_{21}[v_s(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{e}_1) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ &+ \lambda_{22}[v_s(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{e}_2) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ &+ C_a, \end{aligned}$$

where the cost rate $r_s(\boldsymbol{x},\boldsymbol{y})$ and the average cost rate $\bar{r}_s$ are clearly given by

$$\begin{aligned} r_s(\boldsymbol{x},\boldsymbol{y}) &= x_1 + x_2 + y_1 + y_2, \\ \bar{r}_s &= \frac{\alpha_1}{r_1^* - \alpha_1} + \frac{\alpha_2}{r_2^* - \alpha_2}. \end{aligned}$$

In addition, the final term C_a depending on the action a is given by

$$\begin{aligned} C_{10} = \frac{c_0}{\bar{X}}\frac{x_1}{x_1+x_2}[v_s(\boldsymbol{x}-\boldsymbol{e}_1,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ + \frac{kc_0}{\bar{X}}\frac{x_2}{x_1+x_2}[v_s(\boldsymbol{x}-\boldsymbol{e}_2,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})], \end{aligned} \quad (18)$$

$$\begin{aligned} C_{01} = \frac{c_0}{\bar{X}}\frac{y_1}{y_1+y_2}[v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_1) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ + \frac{kc_0}{\bar{X}}\frac{y_2}{y_1+y_2}[v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_2) - v_s(\boldsymbol{x},\boldsymbol{y})], \end{aligned} \quad (19)$$

$$\begin{aligned} C_{11} = \frac{c_1}{\bar{X}}\frac{x_1}{x_1+x_2}[v_s(\boldsymbol{x}-\boldsymbol{e}_1,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ + \frac{kc_1}{\bar{X}}\frac{x_2}{x_1+x_2}[v_s(\boldsymbol{x}-\boldsymbol{e}_2,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ + \frac{c_1}{\bar{X}}\frac{y_1}{y_1+y_2}[v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_1) - v_s(\boldsymbol{x},\boldsymbol{y})] \\ + \frac{kc_1}{\bar{X}}\frac{y_2}{y_1+y_2}[v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_2) - v_s(\boldsymbol{x},\boldsymbol{y})]. \end{aligned} \quad (20)$$

Here $\boldsymbol{e}_1 = (1,0)$ and $\boldsymbol{e}_2 = (0,1)$. It can be easily seen that,

$$\begin{aligned} v_s(\boldsymbol{x}-\boldsymbol{e}_1,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y}) &= -2a_{11}x_1 - 2a_{12}x_2, \\ v_s(\boldsymbol{x}-\boldsymbol{e}_2,\boldsymbol{y}) - v_s(\boldsymbol{x},\boldsymbol{y}) &= -2a_{12}x_1 - 2a_{22}x_2, \\ v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_1) - v_s(\boldsymbol{x},\boldsymbol{y}) &= -2b_{11}y_1 - 2b_{12}y_2, \\ v_s(\boldsymbol{x},\boldsymbol{y}-\boldsymbol{e}_2) - v_s(\boldsymbol{x},\boldsymbol{y}) &= -2b_{12}y_1 - 2b_{22}y_2. \end{aligned}$$

Clearly, the actions 11, 10, and 01 are optimal in (17) if and only if,

$$C_{11} - C_{10} \le 0 \quad C_{11} - C_{01} \le 0, \quad (21)$$
$$C_{11} - C_{10} \ge 0 \quad C_{10} - C_{01} \le 0, \quad (22)$$
$$C_{01} - C_{10} \le 0 \quad C_{11} - C_{01} \ge 0, \quad (23)$$

respectively. With proper substitutions these conditions become

$$\begin{aligned} &\frac{c_1}{c_0 - c_1}\,\frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}} \\ &\ge t(\boldsymbol{x},\boldsymbol{y}) \\ &\ge \frac{c_0 - c_1}{c_1}\,\frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}} \end{aligned} \quad (24)$$

$$t(\boldsymbol{x},\boldsymbol{y}) \ge \frac{c_1}{c_0 - c_1}\,\frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}}, \quad (25a)$$

$$t(\boldsymbol{x},\boldsymbol{y}) \ge \frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}}, \quad (25b)$$

$$t(\boldsymbol{x},\boldsymbol{y}) \le \frac{c_0 - c_1}{c_1}\,\frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}}, \quad (26a)$$

$$t(\boldsymbol{x},\boldsymbol{y}) \le \frac{1-\rho_1^*}{1-\rho_2^*}\,\frac{\frac{(1+k)r_1^*}{\bar{X}} - \lambda_{11} - \lambda_{12}}{\frac{(1+k)r_2^*}{\bar{X}} - \lambda_{21} - \lambda_{22}}, \quad (26b)$$

where

$$t(\boldsymbol{x},\boldsymbol{y}) = \frac{(1+k)(x_1+x_2) - (x_1+kx_2)\rho_1^*}{(1+k)(y_1+y_2) - (y_1+ky_2)\rho_2^*}.$$

In our low interference case, $\frac{c_0 - c_1}{c_1} < 1$. So (25a) implies (25b). Similarly, (26a) implies (26b). Thus, we get switching curves which are a pair of hyperplanes through the origin.

Notice that the switching curves are linear functions of the number of users in different class. This is not at all apparent in the beginning where we observe that the value functions are quadratic function of number of users in different classes. For example, the first condition for the optimality of C_{11} is,

$$\begin{aligned}(c_0 - c_1)(y_1 + y_2)[x_1(a_{11}x_1 + a_{12}x_2) \\ + kx_2(a_{21}x_1 + a_{22}x_2)] \\ - c_1(x_1 + x_2)[y_1(b_{11}y_1 + b_{12}y_2) \\ + ky_2(b_{21}y_1 + b_{22}y_2)] \leq 0,\end{aligned}$$

the left hand side of which is a third degree polynomial in four variables. Surprisingly, the switching curves (24)–(26b) are hyperplanes as they simplify into a linear expression in four variables when the value function coefficient are substituted by their respective values given by (16).

V. Numerical Examples

We now observe numerically the performance of the Improved Dynamic policy compared to the Optimal Static policy, and then compare this improvement with the Both Stations On policy. For the Optimal Static policy, the total average number of users is determined analytically, while for the two other policies we have conducted extensive simulations to estimate the corresponding mean values. The ratio between the total average number of users of the two dynamic policies and that of the Optimal Static policy is plotted in Figures 2–5 for two different scenarios and two different values for the path loss parameter k. In one scenario (called *symmetric user arrivals*), all arrival rates of users are symmetric with $\lambda_{11} = \lambda_{12} = \lambda_{21} = \lambda_{22} = \lambda$ with λ varying from 0 to the stability limit. In the other scenario (called *asymmetric user arrivals*), $\lambda_{12} = \lambda_{21} = 0$ and $\lambda_{11} = \lambda_{22} = \lambda$ with λ again varying from 0 to the stability limit.

For symmetric arrivals in both stations, it is clear that the Improved Dynamic policy is better than the Optimal Static policy as expected, see Figures 2 and 3. The improvement is largest for lower arrival rates, but, with higher arrival rates, it deteriorates clearly. On the other hand, the Both Stations On policy consistently performs better than the Optimal Static policy as well as the Improved Dynamic policy throughout the stability region. There are around 40% less users in the system when the Both Stations On policy is used compared to the Optimal Static policy is used. At higher arrival rates, the improvement is even slightly better.

For asymmetric arrivals, the results are quite different, see Figures 4 and 5. We take an extreme case with only near users in the first station and only far off users in the second station. In such case the Improved Dynamic policy is as good as the Both Stations On policy at lower values of arrival rates. At higher arrival rates, the Both Stations On policy is better but not by very much.

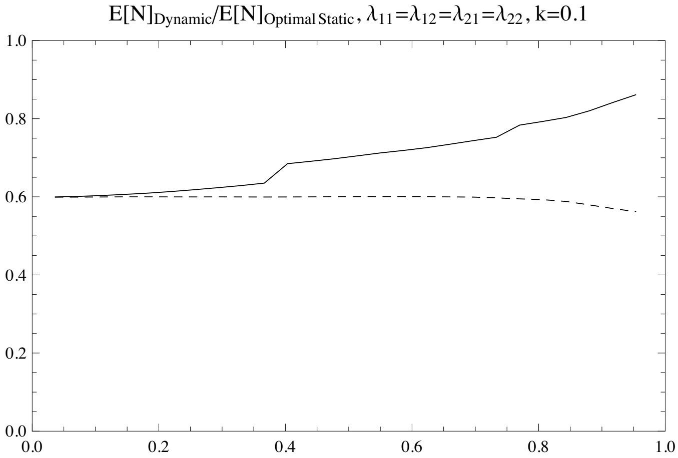

Fig. 2: Symmetric user arrivals with $k = 0.1$. The solid line represents the Improved Dynamic Policy and the dotted line represents the Both Stations On Policy. In the horizontal axis, the arrival rates are plotted normalized to the maximum possible arrival rate that keeps the system stable.

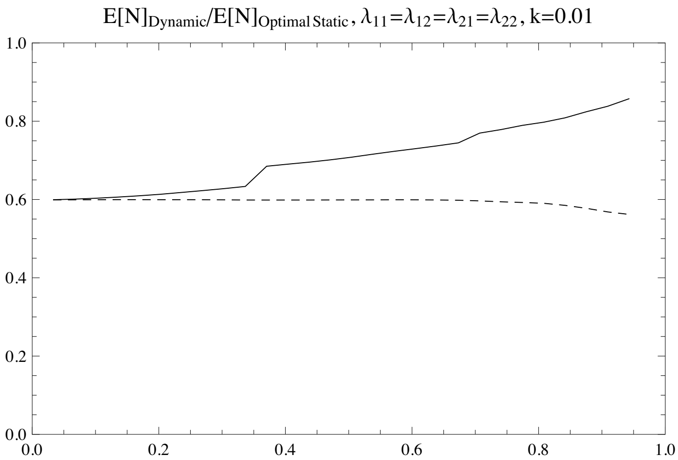

Fig. 3: Symmetric user arrivals with $k = 0.01$. The solid line represents the Improved Dynamic Policy and the dotted line represents the Both Stations On Policy. In the horizontal axis, the arrival rates are plotted normalized to the maximum possible arrival rate that keeps the system stable.

VI. Conclusions

In this paper, we have considered the performance of different resource allocation policies of base stations in a two-cell environment with two classes of users per cell. These policies are chosen on the basis of the knowledge of traffic information of the system, and the one that maintains the overall lowest number of users in

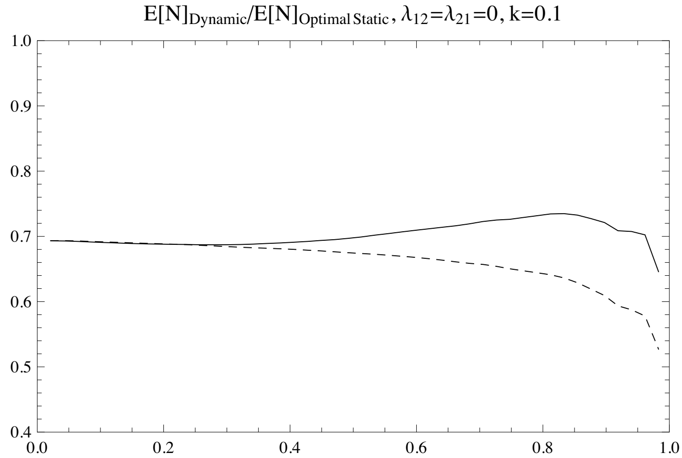

Fig. 4: Asymmetric user arrivals with $k = 0.1$. The solid line represents the Improved Dynamic Policy and the dotted line represents the Both Stations On Policy. In the horizontal axis, the arrival rates are plotted normalized to the maximum possible arrival rate that keeps the system stable.

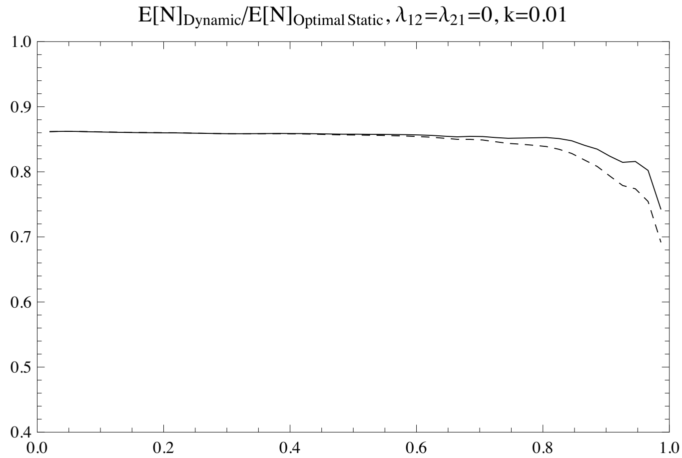

Fig. 5: Asymmetric user arrivals with $k = 0.01$. The solid line represents the Improved Dynamic Policy and the dotted line represents the Both Stations On Policy. In the horizontal axis, the arrival rates are plotted normalized to the maximum possible arrival rate that keeps the system stable.

average is investigated. To this end, the theory of Markov Decision Processes and dynamic programming are used to systematically obtain a policy that improves upon the initial policy.

More specifically, the theory of Markov Decision Processes is utilized for improving upon the Optimal Static policy. Such policy improvement techniques generally converge to optimal or close to optimal policies after a few iteration. In our scenarios, however, such policies are mostly outperformed by the Both Stations On policy indicating that the latter one is indeed close to optimal. In light of previous results which show that similar policies are very strongly optimal when only one class of users are present, our simulation results show this policy performs very well even in a multi-class setting.

Acknowledgement

This research has been partially supported by the HEWINETS (Dynamic Heterogeneous Wireless Access Networks) project, funded by Ericsson, Cassidian Systems, and TEKES.

References

[1] T. Bonald and A. Proutière, “Wireless downlink data channels: user performance and cell dimensioning,” in *Proc. of ACM MobiCom*, Sep. 2003, pp. 339–352.

[2] S. Aalto and P. Lassila, “Impact of size-based scheduling on flow-level performance in wireless downlink data channels,” in *Proc. of ITC-20*, Jun. 2007, pp. 1096–1107.

[3] J. Liu, A. Proutière, Y. Yi, M. Chiang, and H. Poor, “Flow-level stability of data networks with non-convex and time-varying rate regions,” *SIGMETRICS Performance Evaluation Review*, vol. 35, no. 1, pp. 239–250, 2007.

[4] S. Borst, M. Jonckheere, and L. Leskelä, “Stability of parallel queueing systems with coupled service rates,” *Discrete Event Dynamic Systems*, vol. 18, no. 4, pp. 447–472, 2008.

[5] S. Borst, N. Hegde, and A. Proutière, “Interacting queues with server selection and coordinated scheduling – application to cellular data networks,” *Annals of Operations Research*, vol. 170, no. 1, pp. 59–78, 2009.

[6] I. Verloop and R. Núñez-Queija, “Asymptotically optimal parallel resource assignment with interference,” *Queueing Systems*, vol. 65, no. 1, pp. 43–92, 2010.

[7] J. Leino and J. Virtamo, “Determining the moments of queue-length distribution of discriminatory processor-sharing systems with phase-type service requirements,” in *Proc. of NGI*, May 2007, pp. 205–208.

[8] R. A. Howard, *Dynamic Programming and Markov Processes*. New York: Wiley, 1960.

[9] E. Hyytiä, A. Penttinen, and S. Aalto, “Size- and state-aware dispatching problem with queue-specific job sizes,” *European Journal of Operational Research*, vol. 217, no. 2, pp. 357–370, 2012.