

Flow Level Performance Analysis of Wireless Data Networks: A Case Study

Juha Leino, Aleksi Penttinen and Jorma Virtamo
Networking Laboratory, Helsinki University of Technology
P.O.BOX 3000, FIN-02015 HUT, Finland
Email: firstname.lastname@tkk.fi

Abstract—We give an example of flow level performance analysis of data traffic in wireless networks by studying a scenario where two base stations with link adaptation serve in a coordinated fashion downloading users on a road or street between the stations. Due to the dynamic nature of such systems, a detailed flow level analysis is challenging and conventional methods run into computational difficulties. We motivate the detailed analysis by studying the system under different operational goals such as maximum throughput, max-min fairness and balanced fairness, concluding that the performance under these dynamic policies differ significantly from the performance under more tractable static policies. We discuss how the corresponding numerical analyses can be facilitated by applying the notion of balanced fairness and, in particular, introduce a novel approximation method referred to as value extrapolation. Value extrapolation can be applied to approximate any performance measure expressed as the expected value of a random variable which is a function of the system state. The idea of the value extrapolation is to consider the system in the MDP (Markov Decision Processes) setting and to solve the expected value from the Howard equations written for a truncated state space. Instead of a simple truncation, the relative values of states just outside the truncated state space are estimated using a polynomial extrapolation based on the states inside. This leads to a closed system and, unless the system is heavily loaded, allows one to obtain accurate results with remarkably small truncated state spaces.

I. INTRODUCTION

Flow level dynamic performance of a wireless data network and its dependence on traffic is largely beyond reach of mathematical and computational analysis. Whereas the complexity of performance analysis in wireline networks arises typically from the mutual dependencies of interacting users, the peculiarities of wireless domain create additional interdependencies between the users and the system itself. For example, a decision on how bandwidth is allocated to contending users may affect the overall bandwidth available; the available bandwidths are defined by the interference in the network, which itself is typically a function of the bandwidth allocation. Despite its apparent complexity, the importance of analytical performance analysis is undisputed in laying down the foundation for efficient utilization of the scarce wireless resources.

This work was funded by the Academy of Finland, project no. 51355 and Graduate School in Electronics, Telecommunications and Automation (GETA).

We study a scenario where two base stations with link adaptation serve downloading users on a road or street between the stations in a coordinated fashion in order to minimize the adverse effects of interference. Thus we have a dynamic system where file downloads are initiated in a random manner and depart from the system upon completion. The capacity limits of such networks, along with some other cell configurations, have been determined by Bonald et al. [1]. Capacity limit defines the maximum amount of traffic the network can sustain. However, for dimensioning a network more detailed knowledge about the performance, i.e. user perceived quality, is needed. Although the scenario studied in this paper is a coarse simplification of a real system utilizing link adaptation and coordinated operation, it clearly shows that a detailed analysis of the system is required in order to get a reasonable idea of the system performance.

Our contribution is twofold. First, we analyze the two-base-station link adaptation scenario under various performance goals and identify the policies how these goals can be achieved. In contrast to the mere capacity limit, these studies allow finding out the performance also from the user point of view, which is a primary requirement for cost-effective network dimensioning. The difference between static and dynamic policies is also studied.

The second contribution is related to the performance analysis of complex teletraffic systems in general. The traditional approach to modeling a dynamic system usually leads to a multidimensional Markov system which can be analyzed in the equilibrium for a truncated state space by solving a set of linear equations. However, the number of states that can be taken into account this way seriously limits the applicability of the approach.

In this paper we apply two approximate methods that allow computational analysis of significantly larger dynamic systems with good accuracy. Balanced fairness is a resource sharing scheme developed by Bonald and Proutière [2], which makes the dynamic system more tractable. An appealing feature of the BF approach is that its performance is insensitive to traffic details. In other words, the stringent assumptions inherent in the Markov system analysis can be relaxed for the most part without compromising the results. The second computational method of the present paper is a novel approximation scheme referred here to as the value extrapolation. It allows approximating any Markov system performance measure which

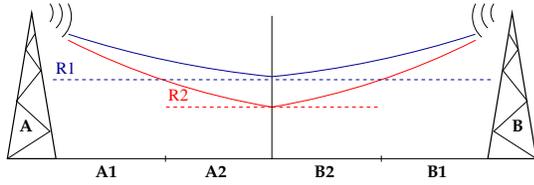


Fig. 1. Example system with two base stations. Areas A1 and A2 are served by the station A and areas B1 and B2 are served by the station B. If both the stations are active simultaneously the maximum rate at A2 or B2 decreases from R_1 to R_2 due to interference.

can be described as an expected value of some state related cost when the system state evolves in time. To reach a comparable approximation accuracy as with the ordinary state space truncation, the method needs to take into account only a fraction of the truncated state space.

The organization of the paper is as follows. Section II describes the model of the system and Section III introduces the methods used to analyze the scenario. In particular, the value extrapolation method is presented and evaluated. Section IV shows the results how the performance of the system is affected by different policies. Section V concludes the paper.

II. EXAMPLE SYSTEM DESCRIPTION

Two base stations, A and B, are used in a co-ordinated fashion to serve elastic traffic, or file downloads, destined to users located on a road between the base stations, cf. Fig. 1. The base station nearest to the user is always used for connection. Link adaptation is modeled as follows. Close to the base stations (in areas A1 and B1) the total downlink rate is always R_1 irrespective of the state of the other station. Further away (in areas A2 and B2) the capacity remains at R_1 only if the other station is not active simultaneously, otherwise the rate decreases to R_2 due to interference.

We describe the system state by the vector $\mathbf{x} = (x_1, x_2, x_3, x_4)$, giving the number of active flows in each area A1, A2, B2, B1, respectively. We use the term flow class interchangeably with the term area. The state space of the system is given by $S = \{x_1, x_2, x_3, x_4 \mid x_i \geq 0, \forall i\}$. In computations we use a truncated state space which is denoted by S' .

As the system state evolves dynamically, we need to fix a policy defining how the network resources are used in any given state. In each state of the system, a policy defines a rate allocation which corresponds to a rate vector $\mathbf{r} = (r_1, r_2, r_3, r_4)$ giving the total rate for each class. The total rate is shared evenly among the flows which belong to a same class by time sharing.

The set of feasible allocations is determined as follows. Let \mathbf{R} be the matrix comprising of column vectors each of which specifies an instantaneous ‘‘operation mode’’ under the constraints described above:

$$\mathbf{R} = \begin{pmatrix} R_1 & 0 & R_1 & 0 & 0 & 0 \\ 0 & R_2 & 0 & R_2 & R_1 & 0 \\ 0 & R_2 & R_2 & 0 & 0 & R_1 \\ R_1 & 0 & 0 & R_1 & 0 & 0 \end{pmatrix}. \quad (1)$$

For example, the first column of \mathbf{R} represents the mode where both base stations serve the flows in the nearest class (areas A1 and B1 are being served), the second column represents the mode where both the stations are active and serve the traffic in the center-most areas. The policy rate vector \mathbf{r} can take the form of any column of \mathbf{R} and, additionally, any convex combination of the columns. These are available through time multiplexing, which is assumed to take place on a fast small time scale compared to flow durations. Thus, the available allocations are defined by the convex hull spanned by the column vectors of \mathbf{R} .

Alternatively, one may determine the hyperplanes that jointly constrain the feasible values of \mathbf{r} . This is just another way to describe the convex hull and can be done using the standard gift wrapping algorithm [3]. The feasible rate vectors \mathbf{r} are constrained by $\mathbf{D}\mathbf{r} \leq \mathbf{e}^T$ (\mathbf{e} is a vector of ones), where, assuming that $\frac{1}{2}R_1 < R_2 < R_1$,

$$\mathbf{D} = \begin{pmatrix} \frac{1}{R_1} & \frac{1}{R_1} & 0 & 0 \\ 0 & 0 & \frac{1}{R_1} & \frac{1}{R_1} \\ \frac{R_2}{R_1^2} & \frac{1}{R_1} & \frac{R_1 - R_2}{R_1 R_2} & \frac{R_1 - R_2}{R_1 R_2} \\ \frac{R_1 - R_2}{R_1^2} & \frac{R_1 - R_2}{R_1 R_2} & \frac{1}{R_1} & \frac{R_2}{R_1^2} \end{pmatrix}.$$

In the case that $0 < R_2 < \frac{1}{2}R_1$ the same matrix applies with the exception

$$d_{33} = d_{42} = \frac{1}{R_1}.$$

III. TOOLS FOR PERFORMANCE ANALYSIS

A. Modeling the system

The first step in analyzing a dynamic system is to determine its capacity limit for a given traffic distribution, i.e. to find the maximum rate the traffic can be served. For the present system this is done in Section IV-A. In practice, the system operation differs from this case in two aspects. First, the performance of the system is measured with respect to user demands instead of the sheer volume of transferred data. Second, one must make resource allocation decisions in a dynamic situation where number of flows in each class is constantly changing and occasionally even zero. Whereas operation at the capacity limit may be easy to describe, more complex performance objectives require state-dependent operating policies. If we describe the system state as stated in the previous section by giving the number of active flows in each class, a wide range of performance objectives becomes available.

In principle, such a setting can be analyzed within a Markovian model, i.e. by assuming that the flows in each class arrive according to a Poisson process and the flow lengths are exponentially distributed. Within the Markov model any operating policy comes down to determining the transition rates downwards from each state, i.e. how the service capacity is shared in each state. For example, if all capacity is allocated to serve traffic in class A2, the corresponding transition rate for class 2 is $\mu_2 R_1$, where $1/\mu_2$ is the mean length of class 2

flows. Having determined the whole transition rate matrix Q , the stationary distribution and performance measures of the system can, in principle, be solved in a straightforward way.

Generally, the basic approach discussed above is extremely inefficient for multidimensional systems. In practice, the infinite state space needs to be truncated and the larger the truncated state space S' is used the more accurate results are achieved. However, determining the steady state distribution for our system so that the state-space is truncated to allow maximum of, e.g., 9 flows in each of the flow classes requires still inversion of a 10000×10000 matrix. Thus, approximate analytical methods are required to obtain more accurate results at higher loads.

In this paper we consider two approximate methods to analyze the system at hand. We approximate the resource sharing policy by a suitable method called balanced fairness that significantly enhances the tractability. On the other hand, we approximate the effect of any given policy on the average performance in the long run by using a novel method called value extrapolation.

B. BF analysis

Balanced fairness is a resource sharing notion that forces the system to be reversible [2]. Reversibility allows deriving the (non-normalized) steady state distribution recursively state-by-state without the need for matrix inversion and we may consider a significantly larger number of states in the state space truncation resulting in higher accuracy. BF can be defined for our system as follows. In state \mathbf{x} , let class i rate be

$$\phi_i(\mathbf{x}) = \frac{\Phi(\mathbf{x} - \mathbf{e}_i)}{\Phi(\mathbf{x})}, \quad i = 1, 2, 3, 4, \quad (2)$$

where $\Phi(\mathbf{x})$ is referred to as the balance function and \mathbf{e}_i denotes a vector of zeros except for the i th component which is 1. Although any positive $\Phi(\mathbf{x})$ makes the system reversible, balanced fairness refers to the unique balance function which utilizes the network resources as efficiently as possible in the sense that in each state at least one resource constraint is saturated. In other words, balance function in state \mathbf{x} is given by

$$\Phi(\mathbf{x}) = \max_i \left\{ (\mathbf{D}\tilde{\Phi})_i \right\}, \quad (3)$$

where $\tilde{\Phi} = (\Phi(\mathbf{x} - \mathbf{e}_1), \dots, \Phi(\mathbf{x} - \mathbf{e}_4))^T$ and \mathbf{D} is the matrix of the hyperplane constraints of the studied system. Thus, the balance function and the corresponding resource sharing policy can be fixed recursively.

The equilibrium distribution is given by

$$\pi(\mathbf{x}) = \frac{1}{G(\boldsymbol{\rho})} \Phi(\mathbf{x}) \rho_1^{x_1} \rho_2^{x_2} \rho_3^{x_3} \rho_4^{x_4}, \quad (4)$$

where G is the normalization constant and ρ_i is the amount of class- i traffic (bit/s).

As $\Phi(\mathbf{x})$ can be derived recursively, we obtain the steady state distribution by going through the states only once, in contrast to the matrix inversion which is required by non-reversible Markov systems. Another nice property of balanced

fairness is its insensitivity. The steady state distribution depends on the traffic characteristics only via traffic load in each class. Using this resource sharing scheme, the performance remains unchanged under very relaxed conditions on the arrival process and flow size distribution [2]. It may also be used as an approximation of utility based sharing schemes such as proportional fairness.

Alternatively, one may define the balance function using the operation mode rate vectors (1). In this case the recursion step takes the form of an LP-problem [5]:

$$\begin{aligned} \Phi(\mathbf{x}) &= \min_{\mathbf{q}} \mathbf{e}^T \mathbf{q}, \\ \mathbf{R}\mathbf{q} &= \tilde{\Phi}, \\ \mathbf{q} &\geq \mathbf{0}, \end{aligned} \quad (5)$$

where \mathbf{e} is a vector of ones.

C. Value extrapolation

The idea of the value extrapolation is to consider the system in the MDP (Markov Decision Processes) setting, cf. e.g. [6], [7], and to solve the expected value of a performance measure from the Howard equations written for a truncated state space. Instead of a simple truncation, the relative values of states just outside the truncated state space are estimated using a polynomial extrapolation based on the states inside. This leads to a closed system and, unless the system is heavily loaded, allows one to obtain accurate results with remarkably small truncated state spaces. Here we give a formal definition of the method.

A policy R specifies a feasible capacity allocation in each state. When a policy R is given, the state transition intensities $q_{\mathbf{x},\mathbf{y}}(R)$ are known. Assume now that the performance measure is described as a revenue rate $r_{\mathbf{x}}(R)$ at state \mathbf{x} and that we are interested in the expected performance of the system as it evolves in time, i.e. the mean revenue rate $r(R)$. Although this measure can be determined using the steady state probabilities as

$$r(R) = \sum_{\mathbf{x}} r_{\mathbf{x}}(R) \pi(\mathbf{x}),$$

the probabilities are computationally tedious to obtain, as discussed above. An alternative characterization of $r(R)$ provides a way to approximate $r(R)$ with significantly higher accuracy.

Let $v_{\mathbf{x}}(R)$ be the relative value of state \mathbf{x} , i.e. the expected difference in cumulative revenue over infinite time horizon when starting from state \mathbf{x} rather than from equilibrium:

$$v_{\mathbf{x}}(R) = \mathbb{E} \left[\int_0^\infty (r_{X(t)}(R) - r(R)) dt \mid X(0) = \mathbf{x} \right],$$

where $X(t)$ is the state process. With a given policy and performance measure, the steady state average revenue rate can be determined by solving the so-called Howard equations [6],

$$r_{\mathbf{x}}(R) - r(R) + \sum_{\mathbf{y} \in S} q_{\mathbf{x},\mathbf{y}}(R) (v_{\mathbf{y}}(R) - v_{\mathbf{x}}(R)) = 0, \quad \forall \mathbf{x}. \quad (6)$$

In the truncated state space there are $|S'|$ equations and $|S'|+1$ variables. The expected state values are fixed only up to an

additive constant, because only the differences $v_y(R) - v_x(R)$ occur in the equations hence we may set, e.g. $v_0(R) = 0$. Note that the mean revenue rate $r(R)$, our performance measure, is one of the unknown variables solved from this group of equations.

The idea of the value extrapolation method is to calculate $r(R)$ in a truncated state space, which essentially means that we assume something on the behavior of the relative values outside the truncated state space. The simplest truncation to some set S' is to set $q_{x,y} = 0 \forall x \in S', y \notin S'$. Regarding to the relative values of the states, this corresponds to setting $v(\dots, N+1, \dots) = v(\dots, N, \dots)$ in the Howard equations, where N is the maximum number of flows in the truncated state space.

The truncation can be done more intelligently if the relative values of the states behave smoothly outside the truncated state space. More accurate results are achieved if the outside values are extrapolated using the values inside the region. First order polynomial extrapolation is $v(\dots, N+1, \dots) = 2v(\dots, N, \dots) - v(\dots, N-1, \dots)$, and the second order extrapolation is $v(\dots, N+1, \dots) = 3v(\dots, N, \dots) - 3v(\dots, N-1, \dots) + v(\dots, N-2, \dots)$.

A strong motivation for this procedure is that the value extrapolation leads to exact results in certain cases. Consider for example an M/M/1-queue, with a policy that allows free entry to the system and with a cost (negative of the revenue) reflecting the total time in the system (which by Little's result is proportional to the mean queue length). The cost rate in a given state is then simply the number of customers in that state, i.e. the state index itself. Let arrival rate be λ , service rate μ and denote $\rho = \frac{\lambda}{\mu}$. Now the Howard equations can be written as

$$i - r + \lambda(v_{i+1} - v_i) + \mu(v_{i-1} - v_i) = 0, \quad \forall i > 0.$$

The equations are clearly solved by

$$r = \frac{\rho}{1-\rho}, \quad v_{i+1} - v_i = \frac{i+1}{\mu-\lambda},$$

from which by setting $v_0 = 0$, we get

$$v_i = \frac{1}{2} \frac{i(i+1)}{\mu-\lambda}.$$

The behavior of the relative value is a simple quadratic polynomial of the state variable. Thus, extrapolating the relative value with the second order extrapolation yields exact value for $r(R)$ no matter how small the truncated space is. It can be reasoned that for any system with cost related to the time in system, the relative values of states are at least asymptotically (i.e. for higher states) quadratic functions of the state occupancy and therefore one can expect the second order extrapolation to work reasonably well.

The advantage of value extrapolation is that even a few states in the truncated state space may be enough to get relatively accurate estimates for the performance measure. The downside is that the Howard equations need to be solved.

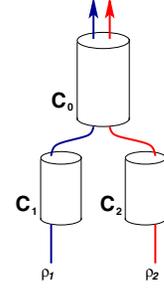


Fig. 2. Example: A tree network.

D. Efficiency of value extrapolation

We demonstrate the efficiency of value extrapolation method in a simple wireline example for which analytical results are available. Consider a two-branch tree network with two flow classes shown in Fig. 2. Flows arrive to both the routes (classes) with the same average intensity. The resources are shared among the classes according to BF. This simple scenario allows deriving exact analytical results under BF [4] and the results are utilized here for comparison purposes.

We are interested in the mean number of flows in the system and illustrate the efficiency of the value extrapolation method in this scenario. Consider the system as a Markov system, where the policy is fixed by BF, i.e. the state transition (service) intensities are given by expression (2) with the corresponding balance function [4]. The truncated state space is $S' = \{x_1, x_2, | x_1, x_2 \leq k\}$ and $r_x(R) = x_1 + x_2$.

Value extrapolation: We extrapolate values $v_{i,k+1}$ and $v_{k+1,i}$ by a suitable method and solve the corresponding Howard equations for $v_{0,0}, \dots, v_{k,k}$ and, in particular, $r(R)$, with $v_{0,0} = 0$. The three methods considered here are:

- The ordinary truncation; states larger than k are ignored, in this case $v_{i,k+1} = v_{i,k}$.
- Linear extrapolation, $v_{i,k+1} = 2v_{i,k} - v_{i,k-1}$.
- Quadratic extrapolation, $v_{i,k+1} = 3v_{i,k} - 3v_{i,k-1} + v_{i,k-2}$.

Fig. 3 shows the comparison of the extrapolation methods in estimating the mean number of flows in the scenario with $C_0 = 1$ and $C_1 = C_2 = 0.6$. It can be seen that the quadratic extrapolation gives quite good approximations even with the truncation parameter $k = 3$. With $k = 6$, the accuracy of the quadratic approximation is excellent up to load 0.8. It slightly overestimates the occupancy but clearly outperforms the simple truncation in this case. However, on higher loads the overestimation becomes significant.

IV. PERFORMANCE EVALUATION

Applying the two methods described in the previous section we can now analyze the wireless example network consisting of the two base stations.

A. Capacity of the system

We describe the capacity of the system by finding the maximum traffic load that the system can sustain. Let x and $1-x$ denote the fraction of traffic in A1 and A2, respectively.

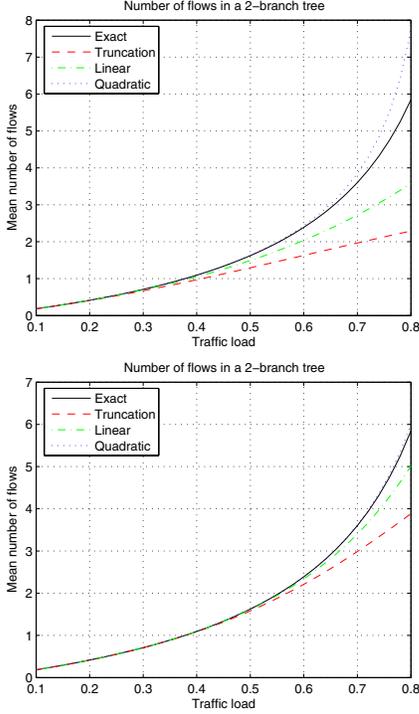


Fig. 3. Value extrapolation in a two-branch tree, $k = 3$ (top) $k = 6$ (bottom)

Assuming symmetry the same fractions hold also for B1 and B2 and the system serves traffic at rate $\mathbf{r} = \frac{R}{2}(x, 1-x, 1-x, x)$, where R is the total capacity to be maximized.

R_{\max} can be obtained by looking at the constraints \mathbf{D} , but can also be derived intuitively as follows. Assume that the load is high and all the classes have traffic.

First, let $0 \leq R_2 < \frac{1}{2}R_1$. Now it is advantageous to serve the traffic in areas A2 and B2 so that only one base station is active. Looking at base station A; it serves area A1 the time x/R_1 , area A2 the fraction of time $(1-x)/R_1$ and, additionally, must remain quiet the fraction of time $(1-x)/R_1$ when station B serves area B2. Due to symmetry

$$R_{\max} = 2 \frac{x + (1-x)}{2(1-x) + x} R_1 = \frac{2R_1}{2-x}.$$

Second, let $\frac{1}{2}R_1 \leq R_2 \leq R_1$. Now the traffic in areas A2 and B2 is served using both the base stations simultaneously. Looking again at station A; it serves class A1 the time x/R_1 , A2 the time $(1-x)/R_2$, but with different rates. Again due to symmetry

$$\begin{aligned} R_{\max} &= 2 \left(\frac{x/R_1}{x/R_1 + (1-x)/R_2} R_1 \right. \\ &\quad \left. + \frac{(1-x)/R_2}{x/R_1 + (1-x)/R_2} R_2 \right) \\ &= \frac{2R_1R_2}{R_1 - R_1x + R_2x}. \end{aligned}$$

B. Performance under different operational policies

We analyze the performance of the system with different capacity allocation policies. We use three different dynamic

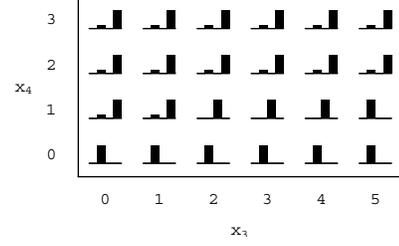


Fig. 4. Example of a nontrivial optimal policy. Capacity allocations for each of the four classes are given by the bar charts for states with different values of x_3 and x_4 but with fixed $x_1 = 0$ and $x_2 = 3$.

policies and compare them to the static policy corresponding to the fixed rates at the capacity limit. The considered policies are:

- The optimal policy maximizing the system throughput
- Max-min fairness
- Balanced fairness

The flow level performance of the policies are compared. We measure the performance of the system with the mean number of active flows which is proportional to the mean flow duration.

The first studied policy optimizes the system performance and can be determined by utilizing the MDP theory. Utilizing the Howard equations (6) and a procedure called policy iteration (see e.g. [6]), the policy minimizing the mean file transfer time can be found. When the optimal policy is known, the value extrapolation method is used to determine the performance. In order to avoid problems caused by the boundary of the state space, the policy iteration is executed using a larger state space than with value extrapolation. The policy iteration is computationally demanding as the Howard equations need to be solved multiple times.

The optimal MDP policy maximizes the mean system throughput. In most states, this is achieved by maximizing the amount of utilized bandwidth but there are some exceptions. An example of nontrivial policy is illustrated in Fig. 4. The capacities are $R_1 = 5$ and $R_2 = 1$ and the traffic intensity of every class is 0.5. For example, while capacity allocation $\mathbf{r} = (0, 1, 0, 5)$ maximizes the throughput in state $(0, 3, 2, 1)$, it is more efficient to serve the flows in region B2 using allocation $\mathbf{r} = (0, 0, 5, 0)$.

While the policy obtained with MDP maximizes the throughput of the system, it does not explicitly consider how the resources are shared among the users. In practise, this may lead to situations where some users efficiently use up all the system capacity while others are left without even acceptable level of service. To avoid such situations, one may impose additional constraints on the resource sharing to guarantee ‘‘fair’’ capacity allocation. We use max-min fairness as an example policy that takes fairness into account. The customers are treated more evenly than with the system optimal policy. In each state, the bandwidth of the flow with least bandwidth is maximized. The performance of the max-min policy is evaluated using value extrapolation.

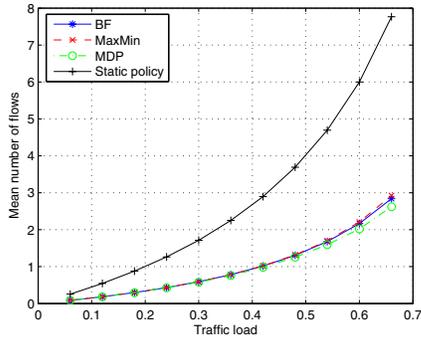


Fig. 5. Mean number of active flows with parameters $R_1 = 5$ and $R_2 = 1$

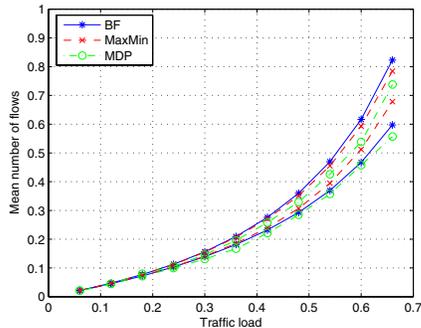


Fig. 6. Mean number of active flows in classes 1 (lower curves) and 2 (upper curves) with parameters $R_1 = 5$ and $R_2 = 1$

The third policy we study is balanced fairness. Balance function and state probabilities are determined recursively using (3) and (4). The mean flow number can then be determined by summing over the state space. BF is significantly faster to evaluate than the value extrapolation method. A significantly larger number of states can be included directly in the analysis to replace the extrapolation need. Thus there is no need to solve the Howard equations.

We illustrate the performance of the policies with parameter values $R_1 = 5$ and $R_2 = 1$. The traffic intensities are assumed equal in all four regions. When value extrapolation method is used, the state space is truncated so that the maximum number of flows in each class is 6. With BF, the corresponding limit is 30. These limits result in accurate results with the traffic loads used.

Fig. 5 illustrates the mean number of active flows with different system loads. The MDP policy has the best performance but the other dynamic policies are almost equal. The static policy of allocating equal bandwidth to all the classes regardless of the system state is significantly worse than the dynamic policies.

The mean number of active flows in traffic classes 1 and 2 are illustrated in Fig. 6. While MDP policy ensures best service for both the classes, the differences between the policies are relatively small. Max-min allocation provides the most equal service to the classes as expected.

The illustrated results are produced with parameter values $R_1 = 5$ and $R_2 = 1$ which fall into region $0 < R_2 < \frac{1}{2}R_1$. In this case, it is beneficial to turn one base station off while the other one serves the flows in the middle region. In addition, we analyzed a scenario with values $R_1 = 5$ and $R_2 = 4$ when it is more efficient to serve the middle classes simultaneously. While the policies differ from the illustrated scenario, the performance is relatively similar.

V. CONCLUSION

Dimensioning of communications networks for elastic data traffic requires knowledge on flow level performance of the system with different traffic loads. This knowledge, however, is difficult to obtain because of the dynamic nature of the system and the flexibility in the system resource usage. An ordinary Markov system model ends up with severe difficulties with state space explosion. It is cumbersome both to build the model and to solve it.

In this work we have analyzed a two-base-station scenario with link adaptation and showed how the practical performance of the system depends on operating policies. Effects of link adaptation and transmission coordination are apparent even in this simplified model. There is a significant gap between the flow level performance resulting from dynamic policies and from static policies. Accordingly, the analysis of the system with static policies is not sufficient to describe the performance for dimensioning purposes.

Irrespective of operational policies the flow level performance of the example system and others alike remains a difficult task. Balanced fairness seems to be a reasonable approximation of the performance. With BF one has to go the state space through only once and no matrix inversions are needed. Furthermore, the results are insensitive to the traffic details which is a very desirable property for dimensioning.

We also contributed to approximate performance analysis by devising an approximation method, the value extrapolation, which can be applied for estimating any performance measure expressed as the expected value over the state space. BF seems to be computationally lighter, but value extrapolation allows analysis of any resource sharing scheme.

These two approaches do not provide an ultimate solution to the problem of state space explosion but they do enable analysis of significantly larger systems than traditional analyses are capable of.

REFERENCES

- [1] T. Bonald, S. Borst and A. Proutière, "Inter-cell scheduling in wireless data networks," in *Proceedings of European Wireless 2005*, 2005.
- [2] T. Bonald and A. Proutière, "Insensitive bandwidth sharing in data networks", *Queueing Systems*, 44, pp. 69–100, 2003.
- [3] S. Skiena, *The Algorithm Design Manual*, Springer-Verlag, 1997.
- [4] T. Bonald, A. Proutière, J. Roberts and J. Virtamo, "Computational aspects of balanced fairness," in *Proc. of the 18th International Teletraffic Congress (ITC-18)*, 2003, pp. 801–810.
- [5] A. Penttinen, J. Virtamo and R. Jäntti, "Performance analysis in multi-hop radio networks with balanced fair resource sharing," *Telecommunications Systems*, in press.
- [6] H. Tijms, *Stochastic Models: An Algorithmic Approach*, John Wiley & Sons, 1994.
- [7] Z. Dziong, *ATM Network Resource Management*, McGraw-Hill, 1997.