Cross-Layer Scheduling of End-to-End Flows Using a Spectrum Server

Chandrasekharan Raman, Roy D. Yates and Narayan B. Mandayam
WINLAB, Rutgers University,
Technology Center of New Jersey
671 Route 1 South,
North Brunswick, NJ 08902.
{chandru, ryates, narayan}@winlab.rutgers.edu

Abstract— **We present a framework for cross-layer scheduling of end-to-end flows in a wireless network with links sharing a common spectrum. Links employ on-off modulation with fixed transmit powers when active. The data rate obtained by an active link is determined by the signal-to-interference ratio on the link. With the knowledge of the link interference gains in the network, the centralized spectrum server schedules the rates on the links and flows on the sessions to maximize a utility function of the source rates. The schedules are a collection of time shared transmission modes (sets of active links).**

I. Introduction

Recent advances in radio technology and spectrum policies have led to a flurry of research in designing spectrum adaptive systems like "cognitive radios" [1], [2]. As a result, radios in future wireless systems are envisaged to be 'smart' and 'interference aware.' Cognitive radios are expected to have the ability to cooperate and dynamically share spectrum with several interfering radios. In addition to the degree of flexibility and adaptability of these radios, the need for global information regarding signals in space, time and frequency plays a prominent role in successful cooperation and coexistence. In that case, efficient open access to spectrum can be aided by impartial "spectrum servers" [3], [4] which can obtain information about the interference environment through measurements contributed by different terminals, and then offer suggestions for efficient coordination among links. In our previous work [5], [6], we consider the role of the spectrum server to derive fair and efficient schedules for throughput maximization of variable rate links. Links employ on/off modulation with fixed power and the rate obtained by each link depends on the interference caused by rest of the links operating simultaneously. We assume that links obtain a non-zero rate for any non-zero SIR. This paper presents a model for cross-layer scheduling of end-to-end flows in a network with links sharing a common spectrum. The centralized spectrum server computes the schedule to maximize the sum of flows through a set of pre-defined routes in the network. We use the cross layer information to maximize the end-to-end session rates in a wireless network.

In the next section, we discuss related work on cross-layer design and scheduling. Section III describes the system model

and presents the optimization problem. We present our results using a simple illustrative example in Section IV and conclude in Section V.

II. Cross-layer Scheduling of End-to-End Flows

End-to-end rate guarantees through link scheduling have been well studied for wired networks [7]. However, rate guarantees in wireless networks are difficult to handle because the shared wireless medium induces a large number of various scheduling constraints. As a result, many cross-layer design approaches have been proposed in the recent literature. Link scheduling in networks have been studied in many contexts, e.g., [8], [9]. The topic of cross-layer scheduling for wireless networks has been addressed in [10], [11].

In [12], the author describes a model in which there exist multiple flows through finite capacity links. The problem is to find a set of optimal rates and flows that maximize a utility function of the source rates subject to constant link capacity constraints. However, our objective is to schedule the links for transmission and find optimal rates and flows subject to rate constraints in the links, due to interference from other links. In [13], the authors propose a fair scheduling algorithm that guarantees end-to-end maxmin fair rates. The scheduling constraints are such that active links at any slot must constitute a matching. Our work provides a schedule which maximizes the sum of utility function of the source rates of origin-destination (OD) pairs, and fair scheduling can be brought out as a special case. The scheduling constraints can be very general and includes schedules consisting of matchings. In [14], the authors describe a cross-layer rate control and scheduling scheme. The rate in a link is a generic function (called the rate-power function) of transmission power of the link. The rates are lower bounded by the sum of source flows in the links. In this work, we provide an explicit characterization for the rate of a link. The rates are variable and depend logarithmically on the Signal-to-Interference Ratio (SIR) of the individual links.

III. System Model

Before we explain the system model, we comment on the notation used in this paper. We use boldface lowercase

characters for vectors and boldface uppercase for matrices. If $\mathbf{a}$ is a vector, $\mathbf{a}^T$ denotes its transpose and $\mathbf{a}^T\mathbf{b} = \sum_i a_i b_i$ represents the inner product of the vectors $\mathbf{a}$ and $\mathbf{b}$. Consider a wireless network with N nodes forming L links sharing a common spectrum. The network can be represented as a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where the nodes in the network are represented by the set of vertices $\mathcal{V}$ of the graph and the logical links are represented by directed set of edges $\mathcal{E}$. Hence the cardinalities $|\mathcal{V}| = N$ and $|\mathcal{E}| = L$. A directed edge from a node m to node n implies that n wishes to communicate data to node m. Let us assume that the network consists of K sessions. A session is specified by an origin-destination (OD) pair. A route r is a sequence of links forming a path in the graph $\mathcal{G}$. We assume that there are R possible routes in the whole network. For a session k, the routes are specified by the $L \times R$ matrix $\mathbf{A}_k$ with entries $\in \{0, 1\}$, where

$$[\mathbf{A}_k]_{lr} = \begin{cases} 1, & \text{if link } l \text{ is a part of route } r, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Let $f_{kj}, \; k = 1, \ldots, K, j = 1, \ldots, R$ be the flow corresponding to the kth session in the jth route. If $\mathcal{R}_l$ denotes the set of routes passing through the link l, we can write the expression for the rate in link l in the session k as

$$r_{lk} = \sum_{j \in \mathcal{R}_l} f_{kj} = \mathbf{a}_{kl}^T \mathbf{f}_k, \quad (2)$$

where $\mathbf{f}_k = [f_{k1} \; f_{k2} \; \cdots \; f_{kR}]^T$ is the vector of flows for the kth session, $k = 1, \ldots, N$, and $\mathbf{a}_{kl}^T$ is the lth row of $\mathbf{A}_k$. Let $\mathbf{r}_k = [r_{1k} \; r_{2k} \; \cdots \; r_{Lk}]^T$, we can then write the link rate vector equation,

$$\mathbf{r}_k = \mathbf{A}_k \mathbf{f}_k. \quad (3)$$

Thus the aggregate rates through links $l = 1, 2, \ldots, L$ are given by

$$\mathbf{r} = \sum_k \mathbf{A}_k \mathbf{f}_k. \quad (4)$$

Each OD pair k in our system gets a rate $y_k = \mathbf{1}^T \mathbf{f}_k, k = 1, \ldots, K$. We are interested in maximizing $\sum_k U_k(y_k)$, the sum of utility functions of the rates in each session.

We now present the interference model for the links in the network. Define the set $\mathcal{T} = \{0, 1, \ldots, M-1\}$ of *transmission modes*, where $M = 2^L$ denotes the number of possible transmission modes. Then the *mode activity vector* $\mathbf{t}_k$ of mode k is a binary vector, indicating the on-off activity of the links. If $\mathbf{t}_k = (t_{1k}, t_{2k}, \ldots, t_{Lk})$ is a mode activity vector, then

$$t_{lk} = \begin{cases} 1, & \text{link } l \text{ is active under transmission mode } k, \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Note that all the M modes may not be valid for transmission. For instance, since the links share a common spectrum, the transmitter and receiver in a node operate in the same channel. Hence the node cannot transmit and receive simultaneously because of self-interference. We refer to this constraint as the *duplexing constraint*. The duplexing constraint implies that

modes corresponding to the adjacent edges in the graph $\mathcal{G}$ are invalid, i.e., the modes should constitute a matching.

We assume a fixed transmitter power on a link l be P_l. If G_{lm} is the link gain from the transmitter of link m to the receiver of link l and σ_l^2 is the noise power at the receiver of link l, the SIR γ_{li} at the receiver of link l in transmission mode i is given by

$$\gamma_{li} = \frac{t_{li} G_{ll} P_l}{\sum_{k \in \mathcal{E}, k \neq l} t_{ki} G_{lk} P_k + \sigma_l^2}. \quad (6)$$

The link gain between a transmitter and receiver takes into account the path loss and attenuation due to shadow fading. We assume that the link gains between each transmitter and receiver are known to the spectrum server. The data rate in each link depends on the SIR in that link. We assume that the transmitter can vary its data rate, possibly through a combination of adaptive modulation and coding. In particular, for a given mode, the transmitter and receiver on a link employ the highest rate that permits reliable communication given the link SIR in that mode. For purposes of this study, we assume that the transmission of other links are treated as Gaussian noise and that a transmission on link l is reliable in a given mode i with a data rate

$$c_{li} = \log(1 + \gamma_{li}). \quad (7)$$

We emphasize here that we do not consider any minimum SIR threshold required at each receiver, i.e., associated with each transmission mode i, a non-zero γ_{li} defines some rate on the link l. Let x_i be the fraction of time that transmission mode i is active. We refer to the vector $\mathbf{x} = [x_0 \; x_1 \; \cdots x_M]^T$ as a schedule without precise specification of sequence of active modes. The average data rate in link l is the time average of the data rates of all the transmission modes that include link l and is given by

$$r_l = \sum_i c_{li} x_i, \quad (8)$$

or in vector form,

$$\mathbf{r} = \mathbf{C}\mathbf{x}, \quad (9)$$

where $\mathbf{C} = [c_{li}]$ is a $L \times M$ matrix with non-negative entries, $\mathbf{r}$ is a real vector of length L and $\mathbf{x}$ is a real vector of length M. Embedded in the matrix $\mathbf{C}$ are the rates obtained in each link l as a part of transmission mode as a result of simultaneous transmissions of multiple links.

In practice, the scheduler will specify a sequence of transmission modes. Typically, this would be done by constructing a frame with N time slots and allocating N_j time slots to each mode j. The fraction of time that mode j is active will be $x_j = N_j/N$. For sufficiently large N, the ratio N_j/N can be made arbitrarily close to any $x_j \in [0, 1]$. In this case, the average rate $\mathbf{r}$ in (9) will represent the average link data rates over one frame. For our analytical model, we optimize these average rates per frame by specification of the time fractions in $\mathbf{x}$, without explicitly specifying the precise slots assigned to

each mode. We denote the set of all feasible schedule vectors by

$$\mathcal{X} = \{\mathbf{x} : \mathbf{1}^T\mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}. \tag{10}$$

The optimization problem for maximizing the sum of utility functions of the rates in each session can be posed as the mathematical program:

$$\max \quad \sum_k U_k(y_k) \tag{11a}$$

$$\text{subject to} \quad y_k = \mathbf{1}^T\mathbf{f}_k, \quad k = 1, \ldots, K, \tag{11b}$$

$$\mathbf{r} = \mathbf{Cx}, \tag{11c}$$

$$\mathbf{r} \geq \sum_k \mathbf{A}_k\mathbf{f}_k, \tag{11d}$$

$$\mathbf{x} \in \mathcal{X}, \tag{11e}$$

$$\mathbf{f}_k \geq 0, \quad k = 1, \ldots, K. \tag{11f}$$

The variables of the above optimization problem are $\mathbf{x}$, $\mathbf{f}_k$, $k = 1, \ldots, K$. If the utility function $U_k(y_k) = y_k$, then (11) maximizes the sum of end-to-end flows of OD pairs. We then get a linear program which can be solved using standard techniques [15]. If $U_k(y_k) = \log(y_k)$, then (11) is a convex optimization problem, which solves for the proportional fair rates [12].

Note that (11c) and (11d) imply that the sum of the flows in each link is upper bounded by a quantity which depends on the schedule. However, in the model described in [12], each link has a finite capacity, which is a constant. The result of the optimization program (11) is a set of transmission modes along with the time fraction of operation of these modes and the flows in each route. Appropriate activity of the modes makes the transport of end-to-end flows possible.

IV. SIMULATION RESULTS

We discuss a simple illustrative example in this section. We consider a network of 5 nodes in a line as shown in Figure 1, each node separated by a distance d from the other node. We label the nodes 1 through 5. These nodes form the vertices $\mathcal{V}$ of a complete graph $\mathcal{G}$. Thus, there are $^5C_2 = 10$ links in the network. Thus each node may be able to transmit to any other in the network in just one hop. Because of the duplexing constraint, the links that transmit in any slot constitute a matching in the complete graph $\mathcal{G}$. Thus, the number is transmission modes in the network is equal to the number of non-trivial matchings in the $\mathcal{G}$, i.e., 25. The effect of interference between the links are captured by the matrix $\mathbf{C}$. The interference gain G_{lj} between the transmitter of link j and the receiver of a link l is given by $G_{lj} = d^{-4}$. The transmit powers are fixed for all transmissions.

We consider a single session originating at node 1 and ending at node 5. Note that there are 8 paths in the network for this OD pair. The objective is to maximize the flow in the network for the given OD pair. Given the distance d between the nodes, we can calculate the SIR for links in every possible mode and then construct the matrix $\mathbf{C}$ for a fixed transmit power. By solving (11), we obtain the routes and the schedule
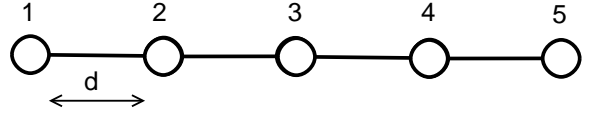

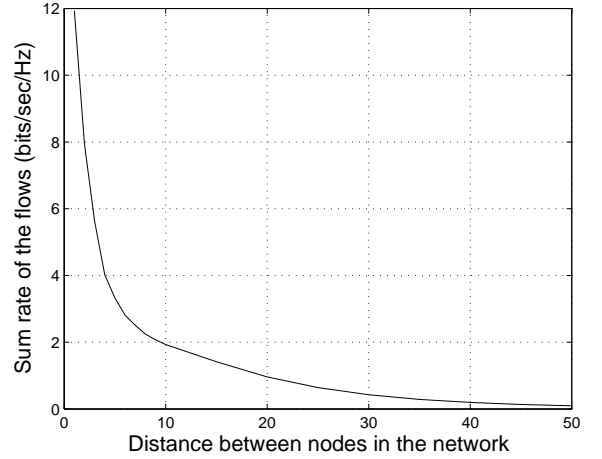Fig. 1. Network with 5 nodes in a line. Each link is of length d

Fig. 2. Variation of source rates with the distance between the nodes for a fixed transmit power. $d = 10$ corresponds to 20 dB received SNR.

for the modes required to obtain the optimal flows in these routes.

Figure 2 shows the variation of sum rate of flows with the distance d between any two nodes in the network. For small values of d, the direct hop is the most optimal route. This also results in the highest sum rate since the mode with the single link $(1, 5)$ can be used. When d increases, there is a four fold increase in the length of direct hop link. Hence the flow between the OD pair decreases rapidly due to the path loss. As d increases further, the single hop link is no more optimal and the flow takes more than one hop to reach the destination. In our example, when $4 < d \leq 6$, two hops are required to maximize the flow. Since the nodes are equally spaced apart, the first hop is at the node 3. The flow decreases with increasing d, but for a two hop case, the link distance increases twice as d. For values of d in the range $6 < d \leq 8$, the flow is maximized when it takes three hops to the destination. For larger values of d, the flow is maximized when the each node transfers to the nearest neighbor and the flow takes 4 hops to the destination. Table I shows the time fraction of activity of the transmission modes for certain values of d. The routes followed by the flows and the scheduled transmission modes are given in the Table II.

V. CONCLUSION

We presented an optimization framework for computing the rates and flow for end-to-end sessions in a wireless network. If the link gains are known, the spectrum server provides the schedule that are a collection of time shared transmission

TABLE I

SCHEDULES FOR THE TRANSMISSION MODES FOR DIFFERENT NODE SEPARATION DISTANCES

Value of d	Flow	Transmission modes used	Fraction of time modes are active
7	2.51	{(1,2)}, {(1,3)}, {(2,4)}, {(3,4)}, {(1,2),(4,5)}	0.0293, 0.1831, 0.3446, 0.1000, 0.3430
7.5	2.375	{(1,2)}, {(1,3)}, {(2,4)}, {(3,4)}, {(1,2),(4,5)}	0.0288, 0.1830, 0.3598, 0.0964, 0.3320
8	2.2441	{(1,2)}, {(1,3)}, {(2,4)}, {(3,4)}, {(1,2),(4,5)}	0.0375, 0.1640, 0.3938, 0.0831, 0.3216
9	2.0735	{(1,2)}, {(2,3)}, {(3,4)}, {(1,2),(4,5)}	0.1138, 0.2856, 0.2856, 0.3125
10	1.9311	{(1,2)}, {(2,3)}, {(3,4)}, {(1,2),(4,5)}	0.1079, 0.2900, 0.2900, 0.3121

TABLE II

ROUTES TAKEN BY THE FLOW FOR DIFFERENT NODE SEPARATION DISTANCES

Range of d	Route (Sequence of links)	Transmission modes used
$0 \leq d \leq 4$	(1,5)	{(1,5)}
$4 < d \leq 6$	(1,3,5)	{(1,3)},{(3,5)}
$6 < d \leq 8$	(1,3,4,5), (1,2,4,5)	{(1,2)}, {(1,3)}, {(2,4)}, {(3,4)}, {(1,2),(4,5)}
$d > 8$	(1,2,3,4,5)	{(1,2)}, {(2,3)}, {(3,4)}, {(1,2),(4,5)}

modes to achieve the end-to-end rates. Rates on the links are limited by the SIR at the receiver of each link. We illustrated the variation of the rates and the routes with the distance between the nodes. Note that this can directly be translated to the variation with respect to received SNRs for a fixed node separation length. A special interesting case is when the optimization program in (11) is solved to obtain fair end-to-end rates for the different sessions in the network. In [5], we show that for a set of randomly placed links that support variable rates, the maxmin fair rates are equal for all links. An interesting question to ask is: what is the maxmin fair rate allocation for the flows in this model? It is not clear as to what constitutes the bottleneck for the links in our model. Another interesting future work is to come up with a distributed scheduling approach, which takes into account only local information of the nodes and links.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201–220, Feb 2005.

[2] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. of Allerton Conf. on Comm., Control and Computing*, Oct 2004.

[3] N. Mandayam, "Cognitive algorithms and architectures for open access to spectrum," *Conf. on the Economics, Technology and Policy of Unlicensed Spectrum, East Lansing, MI*, May 2005. http://quello.msu.edu/conferences/spectrum/program.htm.

[4] M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, and J. Evans, "DIMSUM-Net: New directions in wireless networking using coordinated dynamic spectrum access," in *IEEE WoWMoM05*, June 2005.

[5] C. Raman, R. Yates, and N. Mandayam, "Scheduling variable rate links via a spectrum server," in *Proc. IEEE DySPAN 2005*, 2005. Baltimore, MD.

[6] R. Yates, C. Raman, and N. Mandayam, "Fair and efficient scheduling of variable rate links via a spectrum server," in *Proc. IEEE ICC 2006*, to appear, Jun 2006. Istanbul, Turkey.

[7] A. Parekh and R. Gallager, "A generalized processor sharing approach to flow control - the single node case," *IEEE/ACM Trans. on Networking*, vol. 1, pp. 344–357, Jun 1993.

[8] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *IEEE Trans. Info. Theory*, vol. 34, pp. 910–917, Sept 1988.

[9] L. Tassiulas and A. Ephremides, "Jointly optimal routing and scheduling in packet radio networks," *IEEE Trans. Info. Theory*, vol. 38, pp. 165–168, Jan 1992.

[10] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multi-hop wireless networks," *IEEE Infocom*, 2003.

[11] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: Joint routing, scheduling and power control," in *IEEE Infocom*, 2004.

[12] F. Kelly, "Charging and rate control for elastic traffic," *Euro. Trans. Telecommun.*, vol. 8, pp. 33–37, Jan/Feb 1997.

[13] S. Sarkar and L. Tassiulas, "End-to-end bandwidth guarantees through fair local spectrum share in wireless ad hoc networks," *IEEE Trans. Auto. Control*, vol. 50, pp. 1246–1259, Sept 2005.

[14] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks," in *IEEE Infocom*, 2005.

[15] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.