Simultaneous Routing and Resource Allocation via Dual Decomposition

L. Xiao*, **M. Johansson**† and **S. Boyd***

* Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305
e-mail: {lxiao,boyd}@stanford.edu

† Department of Electrical Engineering and Computer Science
University of California, Berkeley, CA 94720
e-mail: mikaelj@eecs.berkeley.edu

Abstract

In wireless networks, the optimal routing and resource allocation problems are coupled together through link capacities, which influence data routing and are determined by resource (*e.g.*, power and bandwidth) allocation. We formulate the problem of simultaneous routing and resource allocation for wireless data networks as a convex optimization problem and exploit the separable structure of the problem via dual decomposition to develop efficient solution methods.

1 Introduction

Many relevant routing problems for data networks can be formulated as convex multicommodity network flow problems (*e.g.*, [1]), for which many efficient solution methods have been developed (*e.g.*, [2, 3]). The optimal routing tables are highly dependent on the capacities of the communication links, which usually are assumed fixed in the routing literature. The emergence of wireless networks gives new perspectives on optimal routing problems, in particular, when they are coupled with resource allocation problems (*e.g.*, [4, 5]). The capacities of wireless channels are determined by the channel characteristics and communications resources, such as transmit powers, bandwidths, or time-slot fractions allocated to the channels. Adjusting the resource allocation within the network can change the capacities of individual links, influence the optimal routing of data flows and alter the total utility of the network.

In this paper, we consider the simultaneous routing and resource allocation (SRRA) problem for wireless networks. We use a capacitated multicommodity flow model to describe the data flows in the network. Ignoring the detailed transmission protocols and mechanisms, the network flow variables should be interpreted as average flows in bits per second. We assume that the capacity of a wireless link is a concave and increasing function of the communications resources allocated to the link, and the communications resources for groups of links are limited. These assumptions allow us to formulate the SRRA problem as a convex optimization problem over the network flow variables and the communications variables. We exploit the special structure of the SRRA problem via dual decomposition and derive efficient algorithms for solving the dual problem.

2 Network flow model

We use the standard directed-graph model for network topology and a multicommodity flow model for the average behavior of data transmission.

2.1 Network topology

We model the topology of a communication network by a directed graph. In this model, a collection of nodes, labeled by $n = 1, \ldots, N$, can send, receive and relay data through communication links. A communication link is an ordered pair (i, j) of distinct nodes. The presence of a link (i, j) means that the network can support data flow from the start node i to the end node j. We label all the links with integers $l = 1, \ldots, L$. We define $\mathcal{O}(n)$ as the set of links that are outgoing from node n, and $\mathcal{I}(n)$ as the set of links that are incoming to node n.

The network topology can be represented by its node-link incidence matrix. Assume that the network has N nodes and L links, then the entries a_{nl} of the incidence matrix $A \in \mathbf{R}^{N \times L}$ is associated with node n and link l via

$$a_{nl} = \begin{cases} 1, & \text{if } l \in \mathcal{O}(n) \\ -1, & \text{if } l \in \mathcal{I}(n) \\ 0, & \text{otherwise.} \end{cases}$$

Hence, each column of A describes a link, and it has exactly two nonzero entries: one equal to 1 and the other equal to -1, indicating the start and end nodes of the link. Each row of A describes all links incident to a node: the $+1$ entries indicate outgoing links, and the -1 entries indicate incoming links.

2.2 Multicommodity network flows

In the multicommodity flow model, each node can send (different) data to many destinations and receive data from many sources (but we don't consider multicast).

We identify the flows in the network by their destinations, *i.e.*, flows with the same destination are considered as one single commodity regardless of their origins. We label the destination nodes as $d = 1, \ldots, D$, where $D \leq N$ (we can always label the destinations as the first D nodes). For each destination d, we define a source vector $s^{(d)} \in \mathbf{R}^N$, whose entry $s_n^{(d)}$ $(n \neq d)$ denotes the nonnegative flow injected into the network at node n and destined for node d. By the flow conservation law, we define the sink flow at the destination

$$s_d^{(d)} = -\sum_{n, n \neq d} s_n^{(d)}, \tag{1}$$

where the summation is over all n except for $n = d$.

At each intermediate node, flows with the same destination are aggregated and transmitted according to the routing table, which possibly splits the flow among outgoing links. On each link l, we let $x_l^{(d)}$ be the amount of flow destined for node d and call $x^{(d)} \in \mathbf{R}_+^L$ the flow vector with destination d. At each node n, components of the flow vector and the source vector with the same destination satisfy the following flow conservation law:

$$\sum_{l \in \mathcal{O}(n)} x_l^{(d)} - \sum_{l \in \mathcal{I}(n)} x_l^{(d)} = s_n^{(d)}, \quad d = 1, \ldots, D.$$

They can be compactly written as

$$Ax^{(d)} = s^{(d)}, \quad d = 1, \ldots, D, \tag{2}$$

where A is the incidence matrix. Note that (2) implies (1) since $\mathbf{1}^T A = 0$, where $\mathbf{1}$ is the vector of ones.

Finally, we impose the link capacity constraints. Let c_l be the capacity of link l and $t_l = \sum_d x_l^{(d)}$ be the total amount of traffic on link l. We must have $t_l \leq c_l$.

In summary, our network flow model imposes the following group of constraints on the network flow variables $x^{(d)}$, $s^{(d)}$ and t:

$$\begin{array}{l} Ax^{(d)} = s^{(d)}, \quad d = 1, \ldots, D \\ x^{(d)} \succeq 0, \quad d = 1, \ldots, D \\ t_l = \sum_d x_l^{(d)}, \quad l = 1, \ldots, L \\ t_l \leq c_l, \quad l = 1, \ldots, L \end{array} \tag{3}$$

where $\succeq$ means component-wise inequalities. We will use x to denote the collection of flow vectors $x^{(d)}$ and use s to denote the collection of source vectors $s^{(d)}$.

In convex multicommodity network flow problems, the capacities c_l are fixed and one is to minimize a convex cost function of the network variables subject to the constraints (3); see, *e.g.*, [2, 3]. In a wireless network, however, the link capacities c_l are typically functions of the communications resources allocated to the links. These capacity constraints will be described next.

3 Communications model

Now we consider the wireless communication system that supports the data network. In such a system, the capacities of the individual links (channels) depend on the media access scheme and the selection of certain critical parameters, such as transmit powers and bandwidths or time-slot fractions allocated to individual or groups of channels. We refer to these critical communications parameters collectively as *communications variables*, and denote the vector of communications variables by r. We assume that the medium access methods and coding and modulation schemes of the communication system are fixed, but that we can optimize over the communications variables r.

Let r_l be a vector of communications variables associated with link l. In general, the capacity c_l depends not only on r_l, but also on communications variables allocated to other links (due to interferences). However, in this paper we will focus on the case where the link capacity is only a function of local resource allocation r_l, *i.e.*, $c_l = \phi_l(\theta_l)$. For example, communication systems with time-division and frequency-division multiple access (FDMA) fit into this model. We use the following generic model to relate the vector of total traffic t and the vector of communications variables r:

$$\begin{array}{l} t_l \leq c_l = \phi_l(r_l), \quad l = 1, \ldots, L \\ Fr \preceq g, \quad r \succeq 0 \end{array} \tag{4}$$

We make the following assumptions about this model:

- The functions ϕ_l are concave and monotone increasing in r_l. This implies that the first set of constraints are jointly convex in t and r.
- The second set of constraints describe resource limits, such as the total available transmitting power for the links outgoing from the same node.

Capacity formulas of many important communication channel models satisfy the convexity and monotonicity assumptions of the generic model (see, *e.g.*, [6, 7]). Here we will only illustrate how the Gaussian broadcast channel with FDMA fits into this framework.

3.1 Gaussian broadcast channel with FDMA

In this channel model, the transmitters at node n send data to receivers at the end nodes of its outgoing links. The outgoing links $l \in \mathcal{O}(n)$ are assigned disjoint frequency bands with bandwidths $W_l \geq 0$ and powers $P_l \geq 0$. The receivers at the end of the links are subject to independent additive white Gaussian noises with power spectral densities σ_l. The classical Shannon capacity formula (see, *e.g.*, [6]) relates the capacity c_l and the communications variables $r_l = (P_l, W_l)$ by

$$c_l = \phi_l(P_l, W_l) = W_l \log_2 \left(1 + \frac{P_l}{\sigma_l W_l}\right) \tag{5}$$

It can be easily verified that ϕ_l is concave and monotone increasing in the variables (P_l, W_l). So (5) is in the generic form of the first set of constraints in (4).

The communications resource limits are

$$\sum_{l\in\mathcal{O}(i)} P_l \le P_{\text{tot}}^{(n)}, \quad \sum_{l\in\mathcal{O}(n)} W_l \le W_{\text{tot}}^{(n)},$$

which have the generic form of total resource limits (the second set of constraints) in (4).

3.2 The resource allocation problem

In wireless communication systems, many resource allocation problems can be written in the form of maximizing a weighted sum of communication rates (assume that they are concave functions of resources), *i.e.*,

$$\begin{array}{ll} \text{maximize} & \sum_l w_l c_l = \sum_l w_l \phi_l(r_l) \\ \text{subject to} & Fr \preceq g, \quad r \succeq 0 \end{array} \tag{6}$$

where w_l are nonnegative weights. For example, with the Gaussian broadcast channel in section 3.1, we can allocate both the power and bandwidth to maximize the total communication rate.

Many specialized algorithms have been developed for problem (6) by exploiting its structure. For example, if there is only one total resource limit, then it can be solved by the classical water-filling algorithm (*e.g.*, [6]). Actually, water-filling is the one-dimensional version of the dual decomposition method in section 6.

4 The SRRA problem

Combining the network flow model and communications model described in the previous two sections, we now formulate the SRRA problem. We will first present the maximum-utility version of the problem and then discuss many variations of it.

4.1 Maximum-utility SRRA

Consider the operation of a wireless network described by the network flow model (3) and communications model (4). We assume that the utility of each source rate $s_n^{(d)}$ ($n \ne d$) is a concave and increasing function $U_n^{(d)}(\cdot)$. Then the maximum-utility SRRA problem is

$$\begin{array}{ll} \text{maximize} & \sum_d \sum_{n,n\neq d} U_n^{(d)}(s_n^{(d)}) \\ \text{subject to} & Ax^{(d)} = s^{(d)}, \quad d = 1,\dots,D \\ & x^{(d)} \succeq 0, \quad d=1,\dots,D \\ & t_l = \sum_d x_l^{(d)}, \quad l = 1,\dots,L \\ & t_l \le \phi_l(r_l), \quad l=1,\dots,L \\ & Fr \preceq g, \quad r \succeq 0. \end{array} \tag{7}$$

Here the optimization variables are the network flow variables x, s, t and the communications variables r.

This is a convex optimization problem and it can be solved efficiently by, for example, general interior point methods (see, *e.g.*, [8, 9]). Moreover, in the above model, the matrices A and F are sparse and highly structured, and far more efficient algorithms can be developed by exploiting the problem structure.

4.2 Variations of the SRRA problem

Minimum power SRRA. Given the source vectors $s^{(d)}$ to be supported by the network, it is desirable to find the joint routing and resource allocation that minimizes the total transmit power used by the network:

$$\begin{array}{ll} \text{minimize} & w^T r \\ \text{subject to} & \text{the constraints in (7)} \end{array}$$

where $w_i = 1$ if r_i is a power variable and $w_i = 0$ otherwise. Many variations, such as minimizing the maximum power used by any node, or minimizing the total bandwidth required to support the desired traffic, can be handled similarly.

Accounting for delays. One of the most common cost functions in communication network literature is the total delay function $f_{\text{delay}}(x,t) = \sum_l \frac{t_l}{c_l - t_l}$, which is convex if the capacities c_l are fixed. The goal is to minimize this function by selecting the routing variables x, when the source vectors s (the throughput of the network) are given. But this function is not jointly convex in t and r when c_l is substituted by $\phi_l(r_l)$. Another cost function with similar qualitative delay properties is the maximum link utilization (see, *e.g.*, [1])

$$f(t,r) = \max_l \frac{t_l}{\phi_l(r_l)}.$$

This function is quasi-convex, and we can use it as the cost function to be minimized in the SRRA problem.

5 Numerical example

Now we consider a randomly generated wireless network with 50 nodes (see figure 1). The nodes are uniformly distributed in the unit square $[0,1]\times[0,1]$. Two nodes can communicate to each other if their distance is smaller than 0.25 (the graph is strongly connected). The network has 340 links (170 double direction links shown in the figure). We randomly choose five source and destination nodes, labeled $S_1,\dots,S_5$ in figure 1.

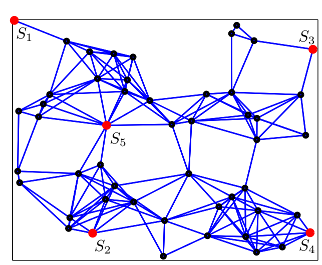
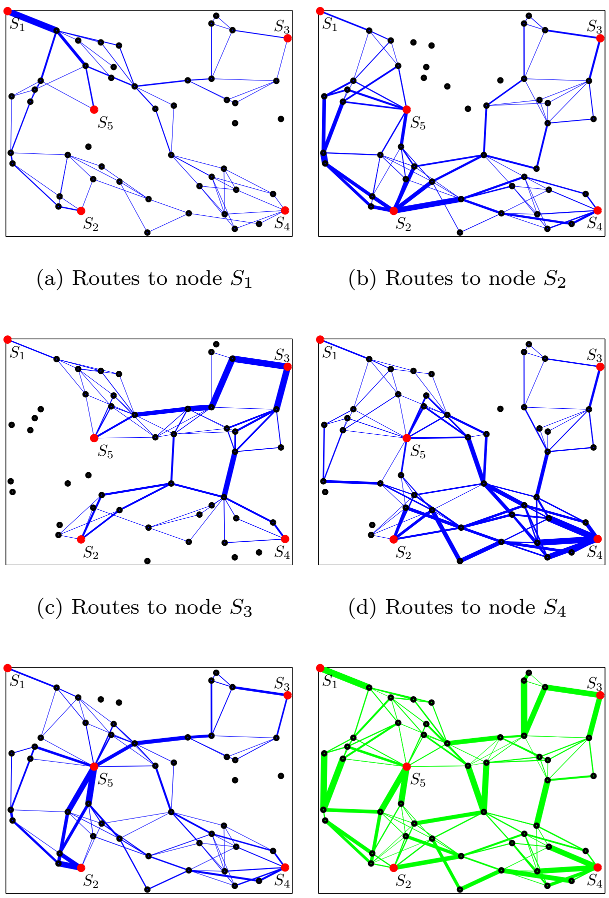


Figure 1: Topology of a randomly generated network.

In this example, we assume that the bandwidth allocation is fixed and there is no interference among links (using FDMA). We consider the joint routing and power allocation problem. Let P_l be the power allocated to link l. Each node has a total power constraint for all its outgoing links, *i.e.*, $\sum_{l\in\mathcal{O}(n)} P_l \le P_{\text{tot}}^n = 100$



(a) Routes to node S_1 (b) Routes to node S_2

(c) Routes to node S_3 (d) Routes to node S_4

(e) Routes to node S_5 (f) Power allocation

Figure 2: Optimal routing and power allocation

for $n = 1, \ldots, N$. Let y_l be the distance between the two end nodes of link l. We use an inverse-square path-loss model: power at the receiver is given by $P_l(y_0/y_l)^2$, where $y_0 = \min_l y_l$ is the reference distance. The noise power σ_l at each receiver is uniformly distributed on $[0.01, 0.1]$. We use the following link capacity function and source utility function for problem (7):

$$\phi_l(P_l) = \log\left(1 + \left(\frac{y_0}{y_l}\right)^2 \frac{P_l}{\sigma_l}\right), \quad U_n^{(d)}(s_n^{(d)}) = \log s_n^{(d)}.$$

We solved the SRRA problem (7) using the dual decomposition method described in section 6. Figure 2(a)–2(e) show the optimal routing tables for the five destinations $S_1, \ldots, S_5$ respectively, and figure 2(f) shows the optimal power allocation. In all these figures, the thickness of each link is roughly proportional to the associated flow variable or the power allocation.

Table 1 shows the source and sink flows which achieve the maximum total utility 17.27. To compare with the SRRA approach, we also solved a maximum utility routing problem with uniform power allocation, where all nodes distribute its total power evenly to its outgoing links, and the results are shown in table 2, which

i	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
1	-3.88	1.11	0.92	1.12	1.13
2	1.03	-16.05	2.93	6.98	6.97
3	0.84	2.69	-9.43	2.69	2.77
4	0.96	4.80	2.46	-18.23	4.80
5	1.05	7.45	3.12	7.44	-15.67

Table 1: Source vectors $s_i^{(d)}$ with SRRA

i	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
1	-2.26	1.03	0.88	1.01	1.37
2	0.56	-13.95	1.73	9.59	5.92
3	0.54	2.07	-6.61	1.97	4.14
4	0.54	6.70	1.55	-16.34	4.20
5	0.62	4.15	2.45	3.77	-15.63

Table 2: $s_i^{(d)}$ with uniform power allocation

give the maximum total utility 12.77. We see that SRRA gives a 35% improvement of performance.

6 The dual decomposition method

In the SRRA problem (7) (the primal problem), the network flow variables x, s, t and the communications variables r are coupled only through the capacity constraints $t_l \le \phi_l(r_l)$. We exploit this almost separable structure via dual decomposition (see, *e.g.*, [10, 9]).

6.1 Formulation of the dual problem

We first form the *partial* Lagrangian, by introducing Lagrange multipliers $p \in \mathbf{R}_+^L$ only for the L coupling constraints $t_l \le \phi_l(r_l)$. This results in

$$\begin{aligned} L(x,s,t,r,p) &= \sum_d \sum_{n, n \ne d} U_n^{(d)}(s_n^{(d)}) - \sum_l p_l(t_l - \phi_l(r_l)) \\ &= \sum_d \left(\sum_{n,n\ne d} U_n^{(d)}(s_n^{(d)}) - \sum_l p_l x_l^{(d)} \right) + \sum_l p_l \phi_l(r_l). \end{aligned}$$

We have eliminated the variables t_l using $t_l = \sum_d x_l^{(d)}$ in the above expression and will write the Lagrangian as $L(x,s,r,p)$ henceforth. The dual function, *i.e.*, the objective function of the dual problem, is defined as

$$V(p) = \sup_{x,s,r} \left\{ L(x,s,r,p) \;\middle|\; \begin{array}{l} Ax^{(d)} = s^{(d)}, \; x^{(d)} \succeq 0 \\ d = 1, \ldots, D \\ Fr \preceq g, \; r \succeq 0 \end{array} \right\}.$$

One immediate observation is that the dual function can be evaluated separately in the network flow variables x, s, t and the communications variables r, *i.e.*,

$$V(p) = V_{\text{net}}(p) + V_{\text{comm}}(p),$$

where $V_{\text{net}}(p)$ can be computed by solving the problem

$$\begin{array}{ll} \text{maximize} & \sum_d \left(\sum_{n,n\ne d} U_n^{(d)}(s_n^{(d)}) - \sum_l p_l x_l^{(d)} \right) \\ \text{subject to} & Ax^{(d)} = s^{(d)}, \quad x^{(d)} \succeq 0, \\ & d = 1, \ldots, D \end{array} \qquad (8)$$

and $V_{\rm comm}(p)$ can be computed by solving the problem

$$\begin{array}{ll} \text{maximize} & \sum_l p_l \phi_l(r_l) \\ \text{subject to} & F \preceq g, \quad r \succeq 0. \end{array} \tag{9}$$

We call (8) the network flow subproblem and (9) the resource allocation subproblem (same as (6)). Both are convex problems. Actually (8) can be completely decomposed into D single-commodity flow problems.

The dual problem associated with the primal (7) is

$$\begin{array}{ll} \text{minimize} & V(p) = V_{\rm net}(p) + V_{\rm comm}(p) \\ \text{subject to} & p \succeq 0. \end{array} \tag{10}$$

This is a convex optimization problem since V is a convex function. Assuming that Slater's condition for constraint qualification holds, strong duality holds (see, *e.g.*, [9, 10]), *i.e.*, the optimal values of the dual problem (10) and the primal problem (7) are equal. Moreover, we can solve the primal problem via the dual.

6.2 Solve SRRA problem via the dual

Notice that the objective function of the primal problem (7) is not strictly convex. So the dual function $V(p)$ is usually non-differentiable (see, *e.g.*, [10]). Effective methods for solving non-differentiable optimization problems include the subgradient methods and cutting plane methods, both of which need to compute $V(p)$ and a subgradient of it at a given $p \succeq 0$.

Compute a subgradient. A subgradient of the convex function V at p is a vector $h \in \mathbf{R}^L$ such that

$$V(q) \geq V(p) + h^T(q-p) \tag{11}$$

for all q. It is a generalization of derivative of differentiable functions (see, *e.g.*, [11]). Given a dual variable $p \succeq 0$, let $x^\star(p), s^\star(p), t^\star(p)$ be one optimal solution to the network flow subproblem (8) and $r^\star(p)$ be one optimal solution to the resource allocation subproblem (9) (they may not be unique). Then a subgradient h of V at p is readily given by (may not be unique either)

$$h_l = \phi_l(r_l^\star(p)) - t_l^\star(p), \quad l = 1, \ldots, L. \tag{12}$$

One can verify this directly from the definition of $V(p)$.

Recover the primal optimal solution. Now suppose that we have numerically solved the dual problem (10), and obtained an optimal dual variable $p^\star$ (see section 6.3 and 6.4). The corresponding solutions $x^\star(p^\star), s^\star(p^\star), t^\star(p^\star)$ and $r^\star(p^\star)$ to the two subproblems (8) and (9), however, may not be primal feasible. In particular, they usually don't not satisfy the capacity constraint $t \preceq \sum_d x^{(d)}$, which was relaxed when forming the dual problem. This is a typical phenomenon for problems with non-strictly convex primal objective functions, where the primal optimal solution is usually a nontrival convex combination of the extreme subproblem solutions (see, *e.g.*, [10] chapter 6). One way to overcome this difficulty is to add a small strictly-convex regularization term to the primal objective function. For example, we added a small quadratic term of x to the utility function in the example of section 5. This approach is closely related to augmented Lagrangian and proximal point methods (*e.g.*, [12]).

Next we discuss two algorithms for solving the dual problem (10): the subgradient methods and the analytic center cutting-plane method (ACCPM).

6.3 The subgradient methods

In subgradient methods, we start with an initial point $p^{(1)}$. At each iteration $k = 1, 2, 3, \ldots$, we compute the dual function $V(p^{(k)})$ and a supergradient $h^{(k)}$ (see Equation (12)), then update the dual variable by

$$p^{(k+1)} = \left[p^{(k)} - \alpha_k h^{(k)}\right]_+ \tag{13}$$

where $[\cdot]_+$ denotes projection onto the nonnegative orthant, and α_k is a positive scalar stepsize. There are many ways to select the stepsizes in subgradient methods. One simple condition for convergence is (*e.g.*,[11])

$$\lim_{k\to 0} \alpha_k = 0, \quad \text{and} \quad \sum_{k=1}^{\infty} \alpha_k = \infty.$$

In the example of section 5, we used the stepsize rule $\alpha_k = \beta/k$. Figure 3 shows the dual objective function versus the iteration number for $\beta = 0.1$ and $\beta = 0.2$.

Notice that the subgradient component h_l can be obtained locally at link l based on its own traffic t_l and available capacity $\phi_l(r_l)$. So the subgradient methods can be implemented distributedly at each link, without a central coordinator. Distributed algorithms based on the subgradient methods have been developed for rate control problems in data networks (*e.g.*, [13, 14]).

For extensive accounts of the subgradient methods, as well as many variations, see, *e.g.*, [11, 10].

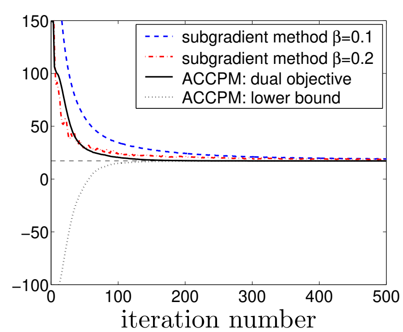


Figure 3: Dual function value versus iteration number

6.4 The analytic center cutting-plane method

ACCPM can be viewed as a localization method where we refine the localization polyhedron at each iteration based on a subgradient computed at its analytic center.

Here we will describe ACCPM in the epigraph space $z = (p, v) \in \mathbf{R}^{L+1}$. Let $\mathcal{P} = \{z \mid Az \preceq b\}$ be a bounded polyhedron, where $A \in \mathbf{R}^{m\times(L+1)}$ and $b \in \mathbf{R}^m$, then its analytic center is defined as

$$\mathbf{acent}(\mathcal{P}) = \arg\max_{z\in\mathcal{P}} \sum_{i=1}^{m} \log\left(b_i - a_i^T z\right)$$

where a_i^T is the ith the row of the matrix A.

We start with a polyhedron $\mathcal{P}^{(0)} = \{z \mid A^{(0)}z \preceq b^{(0)}\}$ which is bounded and contains the optimal solution $z^\star = (p^\star, V^\star)$. For example, it can be a box in $\mathbf{R}^{L+1}$

$$\mathcal{P}^{(0)} = \{z = (p, v) \mid 0 \preceq p \preceq \overline{p},\ \underline{v} \leq v \leq \overline{v}\}$$

where $\underline{v}$ and $\overline{v}$ are known lower and upper bounds for the optimal value $V^\star$, and $\overline{p}$ is a (component-wise) upper bound for $p^\star$. Then ACCPM can be outlined as

given $\mathcal{P}^{(0)}$ and a required tolerance $\epsilon > 0$.
$k := 0$.
repeat

1. Compute the analytic center $z^{(k)} = \mathbf{acent}(\mathcal{P}^{(k)})$, and obtain a lower bound $\underline{v}$ from duality.
2. Let $p^{(k)}$ be the vector of the first L components of $z^{(k)}$, and compute $V(p^{(k)})$ and a supergradient $h^{(k)}$. Then $(p^\star, V^\star)$ must lie in the halfspace
$$\mathcal{H}^{(k)} = \{(p, v) \mid v \geq V(p^{(k)}) + h^{(k)T}(p - p^{(k)})\}.$$
3. Form the polyhedron $\mathcal{P}^{(k+1)} = \mathcal{P}^{(k)} \cap \mathcal{H}^{(k)}$, and update the upper bound $\overline{v} := \min\{\overline{v}, V(p^{(k)})\}$.
4. If $\overline{v} - \underline{v} \leq \epsilon$, quit; else, let $k := k + 1$.

For computational details and convergence analysis of ACCPM, see [15, 9] and references therein.

We applied ACCPM to solve the dual of the SRRA problem in section 5. The dual objective function and the lower bound are also ploted in figure 3. Compared to the subgradient methods, ACCPM usually converges faster. However, in ACCPM, all previous computed subgradients are needed to form the localization set at each iteration, and we need a central coordinator to compute the analytic center.

7 Conclusions

We considered the problem of simultaneous optimal routing and resource allocation for wireless data networks. The network routing problem and the resource allocation problem interact through the capacity constraints of communication links. Using a multicommodity network flow model and assuming that the link capacities are concave functions of the associated resource variables, we formulated the SRRA problem as a convex optimization problem. We exploited the separable structure of the SRRA problem via dual decomposition, and solved the dual problem by the subgradient methods and analytic center cutting-plane method.

Acknowledgement

The authors are grateful to Haitham Hindi and Andrea Goldsmith for helpful discussions. This research was sponsored in part by AFOSR grant F49620-01-1-0365, by DARPA contracts F33615-99-C-3014 and MDA972-02-1-0004, and by MARCO contract 2001-CT-888.

References

[1] D. P. Bertsekas and R. G. Gallager. *Data Networks*. Prentice Hall, 1992.

[2] D. P. Bertsekas. *Network Optimization: Continuous and Discrete Models*. Athena Scientific, 1998.

[3] A. Ouorou, P. Mahey, and J.-Ph. Vial. A survey of algorithms for convex multicommodity flow problems. *Management Science*, 46(1), January 2000.

[4] J. Zander. Radio resource management in future wireless networks: Requirements and limitations. *IEEE Commun. Mag.*, pages 30–36, Aug. 1997.

[5] E. M. Royer and C.-K. Toh. A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Pers. Commun.*, 6:46–55, Apr. 1999.

[6] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons, 1991.

[7] L. Li and A. J. Goldsmith. Capacity and optimal resource allocation for fading broadcast channels: Part I: Ergodic capacity. *IEEE Transactions on Information Theory*, 47(3):1103–1127, March 2001.

[8] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics. SIAM, 1994.

[9] S. P. Boyd and L. Vandenberghe. *EE364 Course reader: Introduction to Convex Optimization with Engineering Applications*. Stanford Univ., 1998.

[10] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, second edition, 1999.

[11] N. Z. Shor. *Minimization Methods for Non-differentiable Functions*. Springer-Verlag, 1985.

[12] R. T. Rockafellar. Augmented lagrangians and applications of the proximal point algorithms in convex programming. *Mathematics of Operations Research*, 1:97–116, 1976.

[13] F. P. Kelly, A. Maulloo, and D. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, 1998.

[14] S. H. Low and D. E. Lapsley. Optimization flow control — I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, December 1999.

[15] J.-L. Goffin and J.-P. Vial. Convex nondifferentiable optimization: A survey focussed on the analytic center cutting plane mathod. Technical Report 99.02, McGill University, Canada, February 1999.