Cost Minimization for Cooperative Traffic Relaying between Primary and Secondary Networks

Feng Tian, *Member, IEEE,* Xu Yuan, *Member, IEEE,* Y. Thomas Hou, *Fellow, IEEE,* Wenjing Lou, *Fellow, IEEE,* and Zhen Yang

Abstract—Cooperation between primary and secondary networks offers significant benefits in data forwarding. But cost implication in such cooperation is not well understood. In this paper, we explore cost incurred in both primary and secondary networks when they are allowed to relay each other's traffic in a cooperative manner. We model costs in both networks and formulate a multiobjective optimization problem. For this problem, we present a novel algorithm to construct an ϵ-approximation curve and prove its error bounds in both cost dimensions. Based on the ϵ-approximation curve, we develop three important applications. The first application is to show the minimum cost value for a single objective (by fixing the other objective as constant) or the relationship between the two objectives over the entire range of possible values. The second application is to address different cost parameters in the primary and secondary networks. We show how to obtain a new approximation curve by scaling the original ϵ-approximation curve with appropriate factors and quantify its error bounds. The third application is to use the ϵ-approximation curve to study a single objective optimization problem with a guaranteed error bound. The results in this paper offer some deep theoretical insights on potential costs incurred in both networks when they are allowed to relay each other's traffic cooperatively.

Index Terms—Cooperative relaying, spectrum sharing, multiobjective optimization, minimum cost curve, primary network, secondary network, ϵ-approximation curve

1 Introduction

Efficient sharing of radio spectrum has been a central focus of the wireless research community for some years [9], [11], [28], [34]. Most research on spectrum sharing has followed the so-called interweave paradigm [10], under which the secondary users' activities do not overlap with the primary users' in time, frequency, or spatial domains. Another implicit assumption under the interweave paradigm is that there is minimal cooperation between the primary and secondary networks on both the data and control planes.

Recently, there is a growing interest in exploring cooperation between primary and secondary networks. For example, in [2], [10], [13], [14], [20], [21], [25], [26], [29], [33], the authors explored the benefits of unilateral cooperation, i.e., to have secondary users help relay traffic for the primary users. To take cooperation one step further, in [31], [32], the authors advocated a *bilateral* cooperation between primary and secondary networks, where the two networks can help relay each other's traffic. Such bilateral cooperation allows to pool together the resources from both networks so that users in each network can access a much richer network resources from the combined network. It was shown in [31], [32] that such bilateral cooperation brings many potential benefits and flexibilities on both the data and control planes that are otherwise not possible. Note that although the two networks are combined into one at the node level, priority or service guarantee to the primary network traffic can still be enforced by implementing appropriate traffic engineering rules.

From the current regulatory perspective, a complete bilateral cooperation between the primary and secondary networks is well ahead of its time. But this should not prevent us from exploring its potential from a research perspective. The need of this research can be well justified by some real-world applications. As an example, for military communications, suppose the underlying spectrum band is assigned to the Navy's network. Suppose the secondary network comes from another branch of the armed forces (e.g., Army). To efficiently relaying traffic, the secondary network may be granted permission to use nodes in the Navy's network to relay its traffic. Likewise, the primary (Navy) network can use the secondary network to help relaying its own traffic. There are many potential benefits for cooperative relay between the primary and secondary networks, which we summarize as follows. From networking perspective, the improved network connectivity, increased flexibility in power control, scheduling and routing all translate into improved forwarding performance for both primary and secondary users' traffic. From spectrum-sharing perspective, the ability to access other network infrastructure helps improve spatial diversity, thus allowing users to tap unused spectrum in the spatial domain. From economic perspective, sharing of two network infrastructures reduces the cost of building each infrastructure separately (by allowing tapping of another network's resource), thus helping enable traditionally underserved population and areas to benefit from

F. Tian, X. Yuan, Y. T. Hou, and W. Lou are with Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. Email: tianfeng1979@gmail.com, {xuy10, thou, wjlou}@vt.edu.
Z. Yang is with Nanjing University of Posts and Telecommunications, Nanjing 210003, China. Email:yangz@njupt.edu.cn.

current and future wireless-enabled services.

To bring the idea of this paradigm one step further, a fundamental understanding of its cost implication is necessary. Obviously, cost has many facets. In this paper, we are interested in the cost incurred when a primary (secondary) node is employed to relay traffic from a secondary (primary) network. Such cost is absent in the interweave paradigm and is only seen when cooperation is involved. There have been some utility/cost related optimization studies on secondary users offering *unilateral* cooperation for the primary network [3], [5], [15], [18], [27], [30]. But for cost involved in bilateral cooperation, there is hardly any result available in the literature.

In this paper, we explore cost incurred in both primary and secondary networks when they are allowed to relay each other's traffic in a cooperative manner. When a secondary (primary) node is used to relay traffic from the primary (secondary) network, a cost would occur (in proportional to primary or secondary traffic's flow rate). We consider the two costs simultaneously by modeling the problem as a multiobjective optimization problem to find the so-called *minimum cost curve*, with each point on the curve representing the minimum cost objective values for the primary and/or secondary network. Unfortunately, minimum cost curve requires to find all (possibly infinite) Pareto-optimal points, which may be infeasible. So we propose to find an ϵ-approximation to the minimum cost curve, which only requires a subset of Pareto-optimal points. We present a novel algorithm to construct an ϵ-approximation curve and prove its approximation errors are less than ϵ in both primary and secondary cost dimensions.

Interestingly, the ϵ-approximation curve has more applications beyond just solving the original multiobjective cost optimization problem. The first important application is that it shows the entire landscape of minimum cost value for a single objective (between the two) or both objectives. The second important application is that it allows to obtain a new approximation curve by simply scaling the original ϵ-approximation curve with appropriate factors. The third important application is that it can serve as an approximate solution to a single objective optimization problem.

The main contributions of this paper are summarized as follows:

- We introduce a cost model for cooperative traffic relaying in primary and secondary networks and formulate a multiobjective optimization problem that finds the *minimum cost curve*, with each point on the curve representing the minimum cost objective values for the primary and/or secondary network. Each point on the curve represents one optimal cooperation solution, i.e., optimal cooperative scheduling, routing and interference management between the primary and secondary networks. Comparing to conventional single objective optimization for only primary or secondary cost, the proposed multiobjective optimization enables primary and secondary networks to flexibly choose their scheduling, routing and interference management to achieve their desired optimal cost objectives. That is, a single objective optimization solution only represents one point on the solution curve of the proposed multiobjective optimization. To our knowledge, this is the first paper that studies cost for cooperative traffic relaying between the primary and secondary networks.
- For the multiobjective optimization problem, we propose a two-phase iterative algorithm to find a subset of Pareto-optimal points and construct the approximation curve. We prove that the constructed approximation curve has approximation errors less than ϵ in both primary and secondary cost dimensions. The solution developed in this paper is generic to the type of multiobjective problems where the two objectives conflict with each other and the problem formulation is in the form of Mixed Integer Linear Program (MILP).
- We identify some important applications for the approximated minimum cost curve. The first application is to show the entire landscape of minimum cost value for a single objective (between the two) or both objectives with approximation performance guarantee. That is, using this curve, for any given cost for the primary (secondary) network, we can immediately find the minimum cost that will incur in the secondary (primary) network with approximation performance guarantee.
- The second application of the approximated minimum cost curve is to derive the new approximated minimum cost curve through simple scaling when the cost parameters change. That is, once we have found the approximated minimum cost curve for a given set of primary and secondary cost parameters, then we can easily find the new approximated minimum cost curve (by scaling with appropriate factors) for a different set of cost parameters. We show that the approximation errors of the newly scaled curve (in both primary and secondary cost dimensions) can be easily found by scaling the approximation errors of the original curve with appropriate factors.
- The third application of the approximated minimum cost curve is that it can be used to solve a single objective optimization problem. We present an algorithm on how to find an approximate solution to a single objective optimization problem based on the approximated minimum cost curve. We also quantify the error bound for the approximate solution to the single objective problem. Our work is the first to show how to apply the approximated minimum cost curve to solve the single objective optimization problem.

The remainder of this paper is organized as follows. Section 2 presents related work. In Section 3, we present cost models for cooperative traffic relaying between primary and secondary networks and formulate a multiobjective optimization problem that minimizes costs in both networks. In Section 4, we present a two-phase iterative algorithm to find a subset of Pareto-optimal points and show how to construct an ϵ-approximation curve. Sections 5 and 6 show how the ϵ-approximation curve can be extended for some important applications. Specifically, in Section 5, we show how to find a new approximate minimum cost curve for a different set of cost parameters based on the original ϵ-approximation

TABLE 1: Notation

Primary Network	
$\hat{\mathcal{N}}_{\mathrm{P}}$	The set of primary nodes
$\hat{\mathcal{L}}_{\mathrm{P}}$	The set of primary sessions
$\hat{f}_{ij}(l)$	The flow rate traversing on link (i,j) that is attributed to primary session $l \in \hat{\mathcal{L}}_{\mathrm{P}}, i,j \in \mathcal{N}$
$\hat{s}(l)$	The source node of primary session $l \in \hat{\mathcal{L}}_{\mathrm{P}}$
$\hat{d}(l)$	The destination node of primary session $l \in \hat{\mathcal{L}}_{\mathrm{P}}$
$\hat{R}(l)$	The minimum data rate requirement of primary session $l \in \hat{\mathcal{L}}_{\mathrm{P}}$
$\hat{\mu}$	The cost parameter for primary users
$\hat{q}_{\mathrm{P}}$	The cost for the primary nodes to relay all secondary flows (also written as $\hat{q}$ when where is no ambiguity)
Secondary Network	
$\mathcal{N}_{\mathrm{S}}$	The set of secondary nodes
$\mathcal{L}_{\mathrm{S}}$	The set of secondary sessions
$f_{ij}(m)$	The flow rate traversing on link (i,j) that is attributed to secondary session $m \in \mathcal{L}_{\mathrm{S}}, i,j \in \mathcal{N}$
$s(m)$	The source node of secondary session $m \in \mathcal{L}_{\mathrm{S}}$
$d(m)$	The destination node of secondary session $m \in \mathcal{L}_{\mathrm{S}}$
$R(m)$	The minimum data rate requirement of secondary session $m \in \mathcal{L}_{\mathrm{S}}$
μ	The cost parameter for secondary users
q_{S}	The cost for the secondary nodes to relay all primary flows (also written as q when where is no ambiguity)
Combined Network	
$\mathcal{N}$	The set of all nodes in the network, $\mathcal{N} = \hat{\mathcal{N}}_{\mathrm{P}} \bigcup \mathcal{N}_{\mathrm{S}}$
C_{ij}	The link capacity of link $(i,j), i,j \in \mathcal{N}$
$x_{ij}[t]$	$= 1$ if node i is transmitting data to node j in time slot t, and is 0 otherwise
$\mathcal{T}_j$	The set of nodes that are located within the transmission range of node $j \in \mathcal{N}$
$\mathcal{I}_j$	The set of nodes that are located within the interference range of node $j \in \mathcal{N}$
T	The number of time slots in a frame
Q_{Id}	The Ideal point
Q_{Nd}	The Nadir point
Q_0	The initial Pareto-optimal point
Q_Z	The final Pareto-optimal point
Q_K	The Pareto-optimal point in the last iteration before the final Pareto-optimal point Q_Z
Q_H	The Pareto-optimal point in the last iteration before the initial Pareto-optimal point Q_0
Q_k	The Pareto-optimal point in k-th iteration, $k = 1, 2, \cdots, K$ for finding the first subset of Pareto-optimal points
Q_{K+j}	The Pareto-optimal point in $(K+j)$-th iteration, $j = 1, 2, \cdots, H$ for finding the second subset of Pareto-optimal points

curve. In Section 6, we show the ϵ-approximation curve can be exploited to solve a single objective optimization problem. Section 7 presents numerical results to validate the ϵ-approximation curve and its various applications. Section 8 concludes this paper.

2 Related Work

Cooperation between primary and secondary networks. There are different ways that primary and secondary networks can cooperate. We focus on cooperative traffic relaying, which is most relevant to this paper. In [2], [10], [13], [14], [20], [21], [25], [26], [29], [33], the authors considered to have secondary users to relay traffic for the primary users. In return, the primary users offer to share (or lease) the spectrum to the secondary users in the time domain [13], [25], [21], [29], [33], frequency domain [2], [14], [21], [26], [33], or spatial domain [13]. In these efforts, cooperation in traffic relaying is unilateral, in the sense that only the secondary network helps to relay traffic from the primary network but not vice versa. On the other hand, the first *bilateral* cooperation for traffic relaying was proposed in [31], where the primary and secondary networks can help relay each other's traffic. As discussed, although the benefits and potential of such a bilateral cooperation were explored in [31], [32], there was no consideration of cost incurred in such cooperation. This fundamental lack of understanding on mutual cost involved in cooperative traffic relaying is the major motivation for our work in this paper.

Cost optimization in primary and secondary networks. Cost has been of concern even under the interweave paradigm [7], [19], [24]. In [7], [19], the authors considered maintenance cost incurred on the secondary networks for channel/link/route switching so as to adapt to the random behavior of the primary users. In [24], the authors considered cost on the secondary users for their cooperative efforts in estimating the distribution of the primary users' activities. Since the primary and secondary networks do not relay each other's traffic under the interweave paradigm, these studies on costs are orthogonal to our work in this paper.

Recently, there have been some utility/cost related optimization studies on secondary users helping relay the primary traffic [3], [5], [15], [18], [27], [30]. In [3], the authors considered cost on the primary users. They designed routing strategies that provide trade-off between primary transmission energy and primary end-to-end session throughput. In [5], [27], [30], the authors considered cost on the secondary users. In [5], the authors modeled the combined primary and secondary networks as a monopoly market using contract theory, where the contract includes spectrum accessing time and relaying power for secondary users. They studied the optimal contract design and proposed an approximate algorithm that achieved a near-optimal contract. In [27], the authors modeled the dynamic interaction between primary and secondary users as a Markov decision process. They applied reinforcement learning algorithm to find near-optimal performance for secondary users while guaranteeing primary users' performance. In [30], the authors modeled the bargaining between the primary and secondary users as a dynamic Bayesian game while assuming that the primary users had incomplete information of secondary transmission energy. They investigated sequential equilibrium under both single-slot and multi-slot bargaining models. In [15], [18], the authors considered cost on both the primary and secondary users. In [15], the authors studied a spectrum leasing model and employed Lyapunov optimization to design a scheduling algorithm to achieve near-optimal performance. In [18], the authors studied multi-hop relay (for primary traffic) in the secondary network through network formation game. The multi-hop relay path was computed via performing the primary player's strategies in the form of link operations. They also devised a distributed dynamic algorithm to obtain a global-path stable network. Note that although some of these prior efforts considered objectives for both primary and secondary users in their papers, the two objectives typically are solved sequentially under separate optimization problems. This is different from the multiobjective formulation that we undertake in this paper.

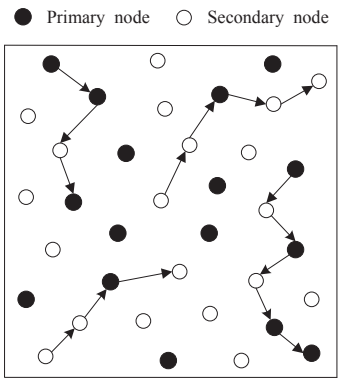


Fig. 1: Cooperation between primary and secondary nodes in traffic relaying.

3 Modeling and Formulation

In this section, we develop a mathematical model for simultaneously minimizing the cost for both primary and secondary networks. Table 1 lists notation in this paper.

Denote $\mathcal{N}$ as the combined set of nodes consisting of both the set of primary nodes $\hat{\mathcal{N}}_{\rm P}$ and the set of secondary nodes $\mathcal{N}_{\rm S}$, i.e., $\mathcal{N} = \hat{\mathcal{N}}_{\rm P} \bigcup \mathcal{N}_{\rm S}$. For simplicity, we follow the protocol interference model (a.k.a. disk model) to represent the impact of transmission and interference ranges. In the combined network (see Fig. 1), denote $\mathcal{T}_j$ as the set of nodes (including both primary and secondary nodes) located within nodes j's transmission range, where j can be either a primary or secondary node (i.e., $j \in \mathcal{N}$). Denote $\mathcal{I}_j$ as the set of nodes (including both primary and secondary nodes) located within a node j's interference range, where j can be either a primary or secondary node. Denote $\hat{\mathcal{L}}_{\rm P}$ and $\mathcal{L}_{\rm S}$ as the set of active primary and secondary sessions, respectively.

3.1 Interference Constraint

Scheduling can be done either in the time domain or frequency domain. In this paper, we consider scheduling in the time domain (with time slots). We assume that a time frame consists of T time slots. Denote $x_{ij}[t]$ as a binary variable to indicate whether node i transmits data to node j in time slot t, i.e.,

$$x_{ij}[t] = \begin{cases} 1 & \text{If node } i \text{ transmits data to node } j \text{ in time slot } t; \\ 0 & \text{otherwise.} \end{cases}$$

For unicast communications, transmit node i can send data to one other node in a time slot, i.e.,

$$\sum_{j}^{j\in\mathcal{T}_i} x_{ij}[t] \le 1 \qquad (i \in \mathcal{N}, 1 \le t \le T)\ . \tag{1}$$

Mutual Interference Constraints. To avoid mutual interference, receive node j can only receive successfully from one active transmit node in a time slot, i.e.,

$$\sum_{k}^{j\in\mathcal{T}_k} x_{kj}[t] \le 1 \qquad (j \in \mathcal{N}, 1 \le t \le T)\ . \tag{2}$$

In general, any receive node $j \in \mathcal{N}$ shall not be interfered by another (unintended) transmit node m $(j \in \mathcal{I}_m)$ whose interference range covers node j in the same time slot, i.e.,

$$x_{ij}[t] + x_{mk}[t] \le 1\ , \tag{3}$$

where $j \in \mathcal{T}_i, j \in \mathcal{I}_m, k \in \mathcal{T}_m, j \in \mathcal{N}, j \ne k, 1 \le t \le T$.

The three constraints in (1), (2) and (3) can be replaced by the following single and equivalent constraint.

$$\sum_{i}^{j\in\mathcal{T}_i} x_{ij}[t] + \sum_{k}^{k\in\mathcal{T}_m} x_{mk}[t] \le 1\ , \tag{4}$$

where $j \in \mathcal{I}_m, j \in \mathcal{N}, j \ne k$ *or* $i \ne m, 1 \le t \le T$.

Self-interference Constraints. We assume half-duplex transceiver is employed at a node. To avoid self-interference, a node i is refrained from transmitting and receiving in the same time slot, i.e.,

$$x_{ij}[t] + x_{ki}[t] \le 1\ \ (i \in \mathcal{N}, j \in \mathcal{T}_i, i \in \mathcal{T}_k, 1 \le t \le T)\ . \tag{5}$$

Again, it can be shown that the three constraints in (1), (2) and (5) can be replaced by the following single and equivalent constraint:

$$\sum_{j}^{j\in\mathcal{T}_i} x_{ij}[t] + \sum_{k}^{i\in\mathcal{T}_k} x_{ki}[t] \le 1 \qquad (i \in \mathcal{N}, 1 \le t \le T)\ . \tag{6}$$

3.2 Rate Requirements and Link Flow Constraints

Primary Rate Requirements. Denote $\hat{f}_{ij}(l)$ as the data rate between nodes i and j that is attributed to primary session $l \in \hat{\mathcal{L}}_{\rm P}$, where $i \in \mathcal{N}$ and $j \in \mathcal{T}_i$. For primary session $l \in \hat{\mathcal{L}}_{\rm P}$, denote $\hat{s}(l)$ and $\hat{d}(l)$ as its source and destination nodes, respectively. For each primary session $l \in \hat{\mathcal{L}}_{\rm P}$, denote $\hat{R}(l)$ as its rate requirement, which is declared by the user. For flexibility and load balancing, we allow flow splitting inside the network. That is, the flow rate of a session may split and merge inside the network in whatever loop-free manner between its source and destination node. As a result, we have the following flow balance constraints at each node that is being traversed by a primary session:

- If node i is the source node of primary session $l \in \hat{\mathcal{L}}_{\rm P}$ (i.e., $i = \hat{s}(l)$), then
$$\sum_{j:j\in\mathcal{T}_i} \hat{f}_{ij}(l) = \hat{R}(l) \qquad (l \in \hat{\mathcal{L}}_{\rm P}, i = \hat{s}(l))\ . \tag{7}$$
- If node i is an intermediate relay node for primary session l (i.e., $i \ne \hat{s}(l)$ and $i \ne \hat{d}(l)$), then
$$\sum_{j:j\in\mathcal{T}_i}^{j\ne\hat{s}(l)} \hat{f}_{ij}(l) = \sum_{k:i\in\mathcal{T}_k}^{k\ne\hat{d}(l)} \hat{f}_{ki}(l)\ \ (l \in \hat{\mathcal{L}}_{\rm P}, i \in \mathcal{N})\ . \tag{8}$$
- If node i is the destination node of primary session l (i.e., $i = \hat{d}(l)$), then
$$\sum_{k}^{i\in\mathcal{T}_k} \hat{f}_{ki}(l) = \hat{R}(l) \qquad (l \in \hat{\mathcal{L}}_{\rm P}, i = \hat{d}(l))\ . \tag{9}$$

Referring to [12], it can be easily verified that if (7) and (8) are satisfied, then (9) must be satisfied. As a result, it is sufficient to list only (7) and (8) in the formulation.

Secondary Rate Requirements. Denote $f_{ij}(m)$ as the data rate between nodes i and j that is attributed to secondary session $m \in \mathcal{L}_{\rm S}$, where $i \in \mathcal{N}$ and $j \in \mathcal{T}_i$. For secondary session $m \in \mathcal{L}_{\rm S}$, denote $s(m)$ and $d(m)$ as the source and destination nodes, respectively. For secondary session $m \in \mathcal{L}_{\rm S}$, denote $R(m)$ as its rate requirement. Then following the same token, we have the following flow balance constraints at each node that is being traversed by a secondary session:

- If node i is the source node of secondary session $m \in \mathcal{L}_{\rm S}$ (i.e., $i = s(m)$), then we have

$$\sum_{j:j\in\mathcal{T}_i} f_{ij}(m) = R(m) \qquad (m\in\mathcal{L}_{\rm S}, i=s(m)) \ . \quad (10)$$

- If node i is an intermediate relay node traversed by secondary session m (i.e., $i\neq s(m)$ and $i\neq d(m)$), then

$$\sum_{j:j\in\mathcal{T}_i}^{j\neq s(m)} f_{ij}(m) = \sum_{k:i\in\mathcal{T}_k}^{k\neq d(m)} f_{ki}(m) \qquad (m\in\mathcal{L}_{\rm S}, i\in\mathcal{N}) \ . \quad (11)$$

- If node i is the destination node of secondary session m (i.e., $i=d(m)$), then

$$\sum_{k}^{i\in\mathcal{T}_k} f_{ki}(m) = R(m) \qquad (m\in\mathcal{L}_{\rm S}, i=d(m)) \ . \quad (12)$$

Just as the case for the primary sessions, it is sufficient to include (10) and (11) as (12) is redundant.

Link Flow Constraints. Denote C_{ij} as the capacity between transmit node i and receive node j. On average (over T time slots), the aggregate flow rates by the primary and secondary sessions on link (i,j) is bounded by the link capacity, we have:

$$\sum_{l\in\hat{\mathcal{L}}_{\rm P}}^{j\neq\hat{s}(l),i\neq\hat{d}(l)} \hat{f}_{ij}(l) + \sum_{m\in\mathcal{L}_{\rm S}}^{j\neq s(m),i\neq d(m)} f_{ij}(m) \leq \frac{1}{T}\sum_{t=1}^{T} C_{ij}\cdot x_{ij}[t] \ . \quad (13)$$

3.3 Cost Model

As discussed in Section 1, cost incurs either when a primary node relays traffic for a secondary session or when a secondary node relays traffic for a primary session. We first consider cost incurred on the primary nodes. Denote $\hat{\mu}$ as the cost at a primary node when it receives and transmits each unit of secondary traffic flow. The total secondary session traffic relayed by primary node i is $\sum_{m\in\mathcal{L}_{\rm S}}\sum_{j:j\in\mathcal{T}_i}^{j\neq s(m)} f_{ij}(m)$. Denote $\hat{q}_{\rm P}$ as the cost over all primary nodes $i\in\hat{\mathcal{N}}_{\rm P}$ for relaying secondary data traffic. We have:

$$\hat{q}_{\rm P} = \hat{\mu}\cdot\sum_{i\in\hat{\mathcal{N}}_{\rm P}}\sum_{m\in\mathcal{L}_{\rm S}}\sum_{j:j\in\mathcal{T}_i}^{j\neq s(m)} f_{ij}(m) \ . \quad (14)$$

Now we consider cost incurred on the secondary nodes. Denote μ as the cost at a secondary node when it receives and transmits each unit of primary traffic flow. The total primary session traffic relayed by secondary node i is $\sum_{l\in\hat{\mathcal{L}}_{\rm P}}\sum_{j:j\in\mathcal{T}_i}^{j\neq\hat{s}(l)}\hat{f}_{ij}(l)$. Denote $q_{\rm S}$ as the cost over all secondary nodes $i\in\mathcal{N}_{\rm S}$ for relaying primary data traffic. We have:

$$q_{\rm S} = \mu\cdot\sum_{i\in\mathcal{N}_{\rm S}}\sum_{l\in\hat{\mathcal{L}}_{\rm P}}\sum_{j:j\in\mathcal{T}_i}^{j\neq\hat{s}(l)}\hat{f}_{ij}(l) \ . \quad (15)$$

3.4 Multiobjective Problem Formulation

In this paper, we are interested in minimizing the costs incurred at *both* the primary and secondary nodes, i.e., a *multiobjective* optimization problem. Putting together the constraints and requirements discussed in this section, we have the following formulation:

BIOPT

$$\begin{aligned}
\min\quad & q_{\rm S} = \mu\cdot\textstyle\sum_{i\in\mathcal{N}_{\rm S}}\sum_{l\in\hat{\mathcal{L}}_{\rm P}}\sum_{j:j\in\mathcal{T}_i}^{j\neq\hat{s}(l)}\hat{f}_{ij}(l)\\
\min\quad & \hat{q}_{\rm P} = \hat{\mu}\cdot\textstyle\sum_{i\in\hat{\mathcal{N}}_{\rm P}}\sum_{m\in\mathcal{L}_{\rm S}}\sum_{j:j\in\mathcal{T}_i}^{j\neq s(m)} f_{ij}(m)\\
\text{s.t.}\quad & \text{Mutual interference constraints: (4);}\\
& \text{Self-interference constraints: (6);}\\
& \text{Primary rate requirements: (7), (8);}\\
& \text{Secondary rate requirements: (10), (11);}\\
& \text{Link flow constraints: (13).}
\end{aligned}$$

In the formulation, $\mu, \hat{\mu}, R(m), \hat{R}(l)$ and C_{ij} are constants, $x_{ij}[t]$ are binary variables, $f_{ij}(m)$ and $\hat{f}_{ij}(l)$ are continuous variables.

In this formulation, satisfying both primary and secondary data rate requirements is listed as two constraints. If these constraints are not satisfied, then the solution is infeasible. Within the feasible solution space, the first objective is to minimize the cost on the secondary network to carry primary users' traffic, while the second objective is to minimize the cost on primary network to carrry secondary users' traffic. The amount of primary users' traffic carried by the secondary network can be controlled by setting cost parameter μ. That is, the smaller μ is, the more primary traffic can be carried by the secondary network. Likewise, the amount of secondary users' traffic carried by the primary network can be controlled by setting cost parameter $\hat{\mu}$. Such tuning of μ and $\hat{\mu}$ is a direct control of the amount of traffic that can relayed between primary and secondary networks. In the extreme case, when we set μ close to 0, meaning that primary user traffic can be relayed almost free by the secondary network, we are encouraging more primary users' traffic to be carried on the secondary network. On the other hand, when we set $\hat{\mu}$ to be extremely large, meaning that it is extremely costly for primary network to relay secondary users' traffic, we are discouraging secondary user's traffic to be carried on the primary network.

BIOPT is in the form of multiobjective optimization problem formulation [6], which differs from traditional single objective optimization problem formulation. For BIOPT, we want to find minimum costs for *both* primary and secondary networks under a set of constraints. This optimization problem is a *multiobjective mixed-integer linear program* (MO-MILP), which is NP-hard in general [6].

For a multiobjective optimization problem (BIOPT), we are pursuing the so-called Pareto-optimal solution. We will give a review of Pareto-optimal solution in Section 3.5. Simply put, under a Pareto-optimal solution, there is no room to further decrease any one objective value. A Pareto-optimal point is the corresponding objective pair $(q,\hat{q})$ of Pareto-optimal solution. Several approaches to solve multiobjective optimization problem can be considered: (i) the weighted sum method [6], [16], [23] , (ii) the Chebyshev norm method [22], [23], and (iii) the ϵ-constraint method [4], [6], [8], [23]. All of them transform a multiobjective problem into a single objective problem, which we discuss separately as follows. In our discussion, ω is denoted as a feasible solution to BIOPT.

- In the weighed sum method, the objective is defined as a nonnegative linear combination of the two objective functions through a parameter $0\leq\beta\leq 1$, i.e., $\min\ \beta q(\omega) + (1-\beta)\hat{q}(\omega)$. Although it is easy to find

one Pareto-optimal point for a given β, it is impossible to find Pareto-optimal points for all β's (infinite number).

- In the Chebyshev norm method, the objective is defined to minimize the minimum value of the weighted Chebyshev norm with weight $0 \leq \beta \leq 1$ over all feasible solutions, i.e., $\min \ \min\{\beta|q_{\rm Id}(\omega)-q(\omega)|, (1-\beta)|\hat{q}_{\rm Id}(\omega)-\hat{q}(\omega)|\}$, where $q_{\rm Id}(\omega)$ and $\hat{q}_{\rm Id}(\omega)$ denote the coordinates of the so-called *Ideal* point [22]. Although the Chebyshev norm method identifies a new β and generates a new Pareto-optimal point in each iteration (instead of enumerating all values blindly like the weighed sum method), the iterations are unable to terminate when there is an infinite number of Pareto-optimal points.
- In the ϵ-constraint method, one objective is chosen as the objective function and the other is transformed into a constraint. The general formulation of ϵ-constraint method can be shown as:

$$\begin{aligned} \min \quad & q(\omega) \\ \text{s.t.} \quad & \hat{q}(\omega) \leq \epsilon \ ; \end{aligned}$$

and

$$\begin{aligned} \min \quad & \hat{q}(\omega) \\ \text{s.t.} \quad & q(\omega) \leq \epsilon \ , \end{aligned}$$

where ϵ denotes an upper bound of the objective function [4]. The ϵ-constraint method determines the range of each objective function and divides it into equal intervals, respectively [4]. For each objective function, it takes one value in the interval as its own upper bound and solves it. In this method, some infeasible solutions can be discarded in order to find the Pareto-optimal points. Although the improved ϵ-constraint methods in [1], [17] can find all discrete Pareto-optimal points to construct an exact Pareto curve, they are limited to solve integer or discrete optimization problems. Their methods cannot be extended to address our BIOPT problem (with both integer and continuous variables) due to the potential existence of infinite and continuous Pareto-optimal points.

3.5 Minimum Cost Curve: Notations and Properties

In this section, we present the necessary notations for a solution to BIOPT. We also present some interesting properties for BIOPT that will help us better understand the problem.

Denote $\mathbf{x}$ as the set of $\{x_{ij}[t], i \in \mathcal{N}, j \in \mathcal{N}, t = 1, 2, \cdots, T\}$, $\mathbf{f}$ as the set of $\{f_{ij}(m), i \in \mathcal{N}, j \in \mathcal{N}, m \in \mathcal{L}_{\rm S}\}$ and $\hat{\mathbf{f}}$ as the set of $\{\hat{f}_{ij}(l), i \in \mathcal{N}, j \in \mathcal{N}, l \in \hat{\mathcal{L}}_{\rm P}\}$. Denote ω as a feasible solution to BIOPT, which consists of $\{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$. Since the cost pair of objectives $q_{\rm S}$ and $\hat{q}_{\rm P}$ in (14)-(15) are both functions of ω, we can also write them as $q(\omega)$ and $\hat{q}(\omega)$ respectively whenever necessary. Note that we omit the subscripts S and P for $q(\omega)$ and $\hat{q}(\omega)$ when there is no confusion.

Denote $(q_1, \hat{q}_1)$ and $(q_2, \hat{q}_2)$ as the objective pairs corresponding to two different feasible solutions ω_1 and ω_2, respectively. We say objective pair $(q_1, \hat{q}_1)$ dominates $(q_2, \hat{q}_2)$ if $q_1 < q_2$ and $\hat{q}_1 \leq \hat{q}_2$, or $q_1 \leq q_2$ and $\hat{q}_1 < \hat{q}_2$. That is, $(q_1, \hat{q}_1)$ dominates $(q_2, \hat{q}_2)$ if solution ω_1 is better than ω_2.

For solution $\omega^\dagger$ with its corresponding objective pair $(q^\dagger, \hat{q}^\dagger)$, if there does not exist another solution with its objective pair $(q, \hat{q})$ satisfying $q < q^\dagger$ and $\hat{q} \leq \hat{q}^\dagger$, or $q \leq q^\dagger$ and $\hat{q} < \hat{q}^\dagger$, then $\omega^\dagger$ is called a *Pareto-optimal solution* to BIOPT and its corresponding objective pair $(q^\dagger, \hat{q}^\dagger)$ is called a *Pareto-optimal point*. In other words, Pareto-optimal solutions are those solutions that have no room to further decrease any one objective value.

Besides Pareto-optimal point, there is also another type of solutions that are also of interest. For solution ω^* with its corresponding objective pair $(q^*, \hat{q}^*)$, if there does not exist another solution with its objective pair $(q, \hat{q})$ satisfying $q < q^*$ and $\hat{q} < \hat{q}^*$, then ω^* is called a *weakly Pareto-optimal solution* and its corresponding objective pair $(q^*, \hat{q}^*)$ is called a *weakly Pareto-optimal point*. Weakly Pareto optimal solutions are those for which decreasing on both objectives *simultaneously* is impossible, but decreasing one objective without increasing the other may be possible. Note that Pareto-optimal points are also *weakly* Pareto-optimal, but weakly Pareto-optimal points are not always Pareto-optimal.

For our problem BIOPT, we are interested in finding the so-called *minimum cost curve*, or *Pareto frontier*, where each point on the curve is a weakly Pareto-optimal point and represents the minimum cost objective values for the primary and secondary networks. What makes a minimum cost curve significant is that it shows the entire landscape of minimum cost value for a single objective (between the two) or both objectives. Once this curve is available, for any given cost for the primary (secondary) network, we can immediately find the minimum cost that will incur in the secondary (primary) network, without the need to solve any optimization problem.

We will discuss how to find the minimum cost curve in the next section. For now, we present some interesting properties. For the first objective function in BIOPT, denote $\hat{\varphi}$ as the primary traffic relayed by the secondary nodes, i.e., $\hat{\varphi} = \sum_{i \in \mathcal{N}_{\rm S}} \sum_{l \in \hat{\mathcal{L}}} \sum_{j:j \in \mathcal{T}_i}^{j \neq \hat{s}(l)} \hat{f}_{ij}(l)$. For the second objective function in BIOPT, denote φ as the secondary traffic relayed by the primary nodes, i.e., $\varphi = \sum_{i \in \hat{\mathcal{N}}_{\rm P}} \sum_{m \in \mathcal{L}} \sum_{j:j \in \mathcal{T}_i}^{j \neq s(m)} f_{ij}(m)$. Hence, we can rewrite the two objective functions in BIOPT as $q_{\rm S} = \mu\hat{\varphi}$ and $\hat{q}_{\rm P} = \hat{\mu}\varphi$, respectively. Denote α_1 and α_2 as two positive scaling factors, i.e., $\alpha_1 > 0$ and $\alpha_2 > 0$. The following property shows how the scaling of cost parameters μ and $\hat{\mu}$ by α_1 and α_2, respectively, will affect the solution in BIOPT.

Property 1: For a weakly Pareto-optimal point $(q^*, \hat{q}^*)$ with a solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ to problem BIOPT, if the cost parameters μ and $\hat{\mu}$ are both scaled by factors of α_1 and α_2, respectively (i.e., $\mu \to \alpha_1\mu$ and $\hat{\mu} \to \alpha_2\hat{\mu}$), then we have a weakly Pareto-optimal point $(\alpha_1 q^*, \alpha_2 \hat{q}^*)$ to the new problem BIOPT (with cost parameters $\alpha_1\mu$ and $\alpha_2\hat{\mu}$). Further, the new solution to $(\alpha_1 q^*, \alpha_2 \hat{q}^*)$ is the same as $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ for the original BIOPT.

A proof of Property 1 is given in Appendix A. This property considers how the point on the minimum cost curve for the original BIOPT is scaled when the cost parameters μ and $\hat{\mu}$ are each scaled by α_1 and α_2, respectively. When we consider the entire minimum cost curve, the following corollary follows.

Corollary 1: Suppose we have the minimum cost curve for problem BIOPT$(\mu, \hat{\mu})$, when the cost parameters μ and $\hat{\mu}$ are each scaled by α_1 and α_2, respectively (i.e., $\mu \to \alpha_1\mu$

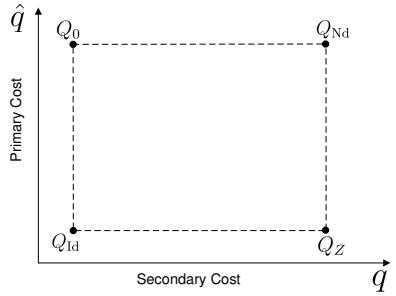


Fig. 2: Finding two Pareto-optimal endpoints Q_0 and Q_Z based on Ideal and Nadir points $Q_{\rm Id}$ and $Q_{\rm Nd}$.

and $\hat{\mu} \rightarrow \alpha_2\hat{\mu}$), the new minimum cost curve for problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ can be found by scaling each point $(q^*, \hat{q}^*)$ on the original minimum cost curve to $(\alpha_1 q^*, \alpha_2\hat{q}^*)$.

4 Approximation Curve

4.1 Roadmap

Given the difficulty in finding all the Pareto-optimal points that are necessary to construct the exact minimum cost curve, we will develop a novel algorithm motivated by the ϵ-constraint method to find an *approximation* curve with performance guarantee (i.e., each point with an approximation error less than ϵ). That is, for any given cost for the secondary (primary) network, the difference in cost for the primary (secondary) networks between the approximation curve and the minimum cost curve is less than ϵ. In this section, we will find such an approximation curve to the minimum cost curve. In our two-phase iterative algorithm, we find a subset of Pareto-optimal points iteratively from BIOPT$(\mu, \hat{\mu})$ and use them to construct an ϵ-approximation curve. The details of each step are given in Sections 4.2 to 4.3. In Section 4.4, we prove that the connected curve is indeed an ϵ-approximation to the minimum cost curve.

4.2 Finding Two Endpoints

As the first step of our algorithm, we show how to find the two Pareto-optimal endpoints, which will determine the range for all the other Pareto-optimal points. To find the two Pareto-optimal endpoints, we need to determine the lower and upper bounds for the Pareto-optimal objective values. This can be done through the so-called *Ideal* and *Nadir* points [6]. Denote $Q_{\rm Id}$ as the *Ideal* point with coordinates $(q_{\rm Id}, \hat{q}_{\rm Id})$, where $q_{\rm Id}$ is obtained by solving $\min q_{\rm S}$ in BIOPT$(\mu, \hat{\mu})$ without minimizing $\hat{q}_{\rm P}$ while $\hat{q}_{\rm Id}$ is obtained by solving $\min \hat{q}_{\rm P}$ without minimizing $q_{\rm S}$. That is, the two objectives for the Ideal point are obtained by solving a single objective optimization problem without considering optimizing the other objective. Obviously, the Ideal point is not achievable, as $q_{\rm Id}$ and $\hat{q}_{\rm Id}$ cannot occur at the same time, due to the conflicting nature between the two objectives in BIOPT$(\mu, \hat{\mu})$. So the Ideal point is only used as a reference point in our algorithm.

Denote $Q_{\rm Nd}$ as the *Nadir* point with coordinates $(q_{\rm Nd}, \hat{q}_{\rm Nd})$, where $q_{\rm Nd}$ is obtained by solving $\min q_{\rm S}$ in BIOPT$(\mu, \hat{\mu})$ by replacing minimizing the other objective $\hat{q}_{\rm P}$ with the assignment $\hat{q}_{\rm P} = \hat{q}_{\rm Id}$ while $\hat{q}_{\rm Nd}$ is obtained by solving $\min \hat{q}_{\rm P}$ in BIOPT$(\mu, \hat{\mu})$ by replacing minimizing the other objective $q_{\rm S}$ with the assignment $q_{\rm S} = q_{\rm Id}$, as shown in Fig. 2. Although the two coordinates in a Nadir point, $q_{\rm Nd}$ and $\hat{q}_{\rm Nd}$, are obtained through solving two separate single objective optimization problems, each of them is obtained by assuming worst case in the other objective value. Since the two objectives go in opposite directions, $q_{\rm Nd}$ and $\hat{q}_{\rm Nd}$ can be considered as the upper bounds for two objectives, respectively. Note that unlike the Ideal point, Nadir point is achievable. Once we have the Ideal point $Q_{\rm Id}(q_{\rm Id}, \hat{q}_{\rm Id})$ and the Nadir point $Q_{\rm Nd}(q_{\rm Nd}, \hat{q}_{\rm Nd})$, we can obtain the two Pareto-optimal endpoints, which we denote as Q_0 and Q_Z, respectively (see Fig. 2). Specifically, we define point Q_0 with its coordinates being $(q_{\rm Id}, \hat{q}_{\rm Nd})$ (i.e., $(q_0, \hat{q}_0)$) and Q_Z with coordinates $(q_{\rm Nd}, \hat{q}_{\rm Id})$.

Lemma 1: Both points $Q_0(q_{\rm Id}, \hat{q}_{\rm Nd})$ and $Q_Z(q_{\rm Nd}, \hat{q}_{\rm Id})$ are Pareto-optimal points for problem BIOPT$(\mu, \hat{\mu})$.

A proof of Lemma 1 is given in Appendix B.

The next lemma shows that any Pareto-optimal point for problem BIOPT$(\mu, \hat{\mu})$ falls between Q_0 and Q_Z.

Lemma 2: For any Pareto-optimal point $Q_k(q_k^\dagger, \hat{q}_k^\dagger)$ for problem BIOPT$(\mu, \hat{\mu})$, we have $q_{\rm Id} \le q_k^\dagger \le q_{\rm Nd}$ and $\hat{q}_{\rm Id} \le \hat{q}_k^\dagger \le \hat{q}_{\rm Nd}$.

A proof of Lemma 2 is given in Appendix C.

4.3 Finding New Pareto-optimal Points

As we discussed in Section 4.1, it is infeasible to find all Pareto-optimal points. In this section, we show how to properly find a subset of Pareto-optimal points between Q_0 and Q_Z. In Phase I, we iteratively reduce the second optimal objective value in BIOPT$(\mu, \hat{\mu})$ to find a subset of Pareto-optimal points from Q_0 to Q_Z. In Phase II, we iteratively reduce the first optimal objective value in BIOPT$(\mu, \hat{\mu})$ to find another subset of Pareto-optimal points from Q_Z to Q_0. Then, we merge the two subsets of Pareto-optimal points into one set that we can use to construct the approximation curve.

Phase I: Finding a Subset of Pareto-optimal Points. In this phase, we show how to find a subset of Pareto-optimal points from Q_0 to Q_Z. Denote ϵ as a small positive constant. We now show how to find Q_1 (a new Pareto-optimal point) based on $Q_0(q_{\rm Id}, \hat{q}_{\rm Nd})$ and ϵ (see Fig. 3(a)). Denote the coordinates of Q_1 as $(q_1^\dagger, \hat{q}_1^\dagger)$. From $Q_0(q_{\rm Id}, \hat{q}_{\rm Nd})$, we reduce the second optimal objective value of Q_0, which is $\hat{q}_{\rm Nd}$, by ϵ to $(\hat{q}_{\rm Nd} - \epsilon)$. That is, we set $\hat{q}_{\rm P} \le \hat{q}_{\rm Nd} - \epsilon$. By replacing the second objective function $(\min \hat{q}_{\rm P})$ with the new constraint $\hat{q}_{\rm P} \le \hat{q}_{\rm Nd} - \epsilon$ in BIOPT$(\mu, \hat{\mu})$, we have the following single objective optimization problem:

$$\begin{aligned} \min \quad & q(\omega) \\ \text{s.t.} \quad & \hat{q}(\omega) \le \hat{q}_{\rm Nd} - \epsilon; \\ & \omega = \{\mathbf{x}, \hat{\mathbf{f}}, \mathbf{f}\}; \\ & \text{Constraints (4),(6)-(8),(10),(11) and (13).} \end{aligned}$$

This optimization problem is in the form of mixed-integer linear program (MILP). In our algorithm, each iteration needs to solve one MILP problem as in Phase 1, which is NP-hard and its computation complexity is exponential. Since the computation complexity is dominated by solving the MILP problem, it is also exponential complexity. Fortunately, all integer variables in this MILP are binary. For binary variables that can only take 0 and 1, a branch-and-cut based solution

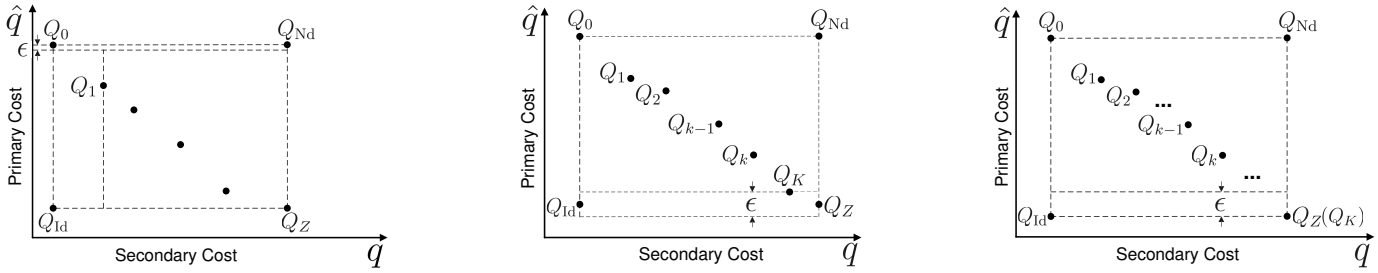


Fig. 3: (a) Finding Pareto-optimal point Q_1 based on Q_0 and ϵ. (b) Pareto-optimal point Q_K is above the final Pareto-optimal point Q_Z. (c) Pareto-optimal point Q_K coincides with the final Pareto-optimal point Q_Z.

procedure used by a commercial solver such as CPLEX is very efficient. Therefore, we use CPLEX to solve all our binary MILP problems, which turns out to be very successful for all practical purposes.

The optimal objective value from the above problem gives $q_1^\dagger$. With $q_1^\dagger$ being fixed, we can now find $\hat{q}_1^\dagger$. This is done by replacing the first objective function ($\min q_{\rm S}$) with the constraint $q_{\rm S} = q_1^\dagger$ in BIOPT($\mu, \hat{\mu}$) and solve the following single objective optimization problem:

$$\begin{aligned} \min \quad & \hat{q}(\omega) \\ \text{s.t.} \quad & q(\omega) = q_1^\dagger; \\ & \omega = \{\mathbf{x}, \hat{\mathbf{f}}, \mathbf{f}\}; \\ & \text{Constraints (4),(6)-(8),(10),(11) and (13).} \end{aligned}$$

The optimal objective value from the above optimization problem gives $\hat{q}_1^\dagger$.

Two remarks on $Q_1(q_1^\dagger, \hat{q}_1^\dagger)$ are in order. First, the new point $Q_1(q_1^\dagger, \hat{q}_1^\dagger)$ is a Pareto-optimal point for problem BIOPT($\mu, \hat{\mu}$), which will be stated in Lemma 3 later in this section. Second, although we start by reducing $\hat{q}_{\rm Nd}$ by ϵ in the first optimization problem, the final gap between $\hat{q}_{\rm Nd}$ and $\hat{q}_1$ can be larger than ϵ, as we show in Fig. 3(a), which will be stated in Property 2.

Denote the process for finding Q_1 (based on Q_0 and ϵ) as the first iteration. We can find points Q_2, Q_3, $\cdots$, Q_K iteratively following the same process, where K is the number of iterations until $\hat{q}_K^\dagger - \epsilon < \hat{q}_{\rm Id}$ (see Fig. 3(b)). In the special case when $\hat{q}_K^\dagger = \hat{q}_{\rm Id}$, Q_K coincides exactly with Q_Z, as shown in Fig. 3(c). Denote $\{Q_0, Q_1, Q_2, \cdots, Q_K, Q_Z\}$ as the Pareto-optimal point subset $\mathcal{H}_1$ and the coordinates of Q_k as $(q_k^\dagger, \hat{q}_k^\dagger)$, $k = 1, 2, \cdots, K$. More formally, we can find Q_k based on Q_{k-1} and ϵ as follows. Given the two optimal objectives $(q_{k-1}, \hat{q}_{k-1})$ for Q_{k-1}, we can reduce the second optimal objective value of $\hat{q}_{k-1}^\dagger$ by ϵ to $(\hat{q}_{k-1}^\dagger - \epsilon)$. That is, we can replace the second objective function ($\min \hat{q}_{\rm P}$) with the constraint $\hat{q}_{\rm P} \le \hat{q}_{k-1} - \epsilon$ in BIOPT($\mu, \hat{\mu}$) and have a single objective optimization problem as following:

$$\begin{aligned} &\text{OPT}(q_{\rm S}) \\ &\begin{aligned} \min \quad & q(\omega) \\ \text{s.t.} \quad & \hat{q}(\omega) \le \hat{q}_{k-1}^\dagger - \epsilon; \\ & \omega = \{\mathbf{x}, \hat{\mathbf{f}}, \mathbf{f}\}; \\ & \text{Constraints (4),(6)-(8),(10),(11) and (13).} \end{aligned} \end{aligned}$$

The optimal objective value of the above problem gives $q_k^\dagger$, the first optimal objective value for Q_k. Likewise, with $q_k^\dagger$ being fixed, we can now find the second optimal objective value $\hat{q}_k^\dagger$ for Q_k. This is done by replacing the first objective function ($\min q_{\rm S}$) with the constraint $q_{\rm S} = q_k^\dagger$ in BIOPT($\mu, \hat{\mu}$), and solving the following single objective optimization problem:

$$\begin{aligned} &\text{OPT}(\hat{q}_{\rm P}) \\ &\begin{aligned} \min \quad & \hat{q}(\omega) \\ \text{s.t.} \quad & q(\omega) = q_k^\dagger; \\ & \omega = \{\mathbf{x}, \hat{\mathbf{f}}, \mathbf{f}\}; \\ & \text{Constraints (4),(6)-(8),(10),(11) and (13).} \end{aligned} \end{aligned}$$

The optimal objective value of the above problem gives $\hat{q}_k^\dagger$.

Such iteration continues until for some k such that $\hat{q}_k^\dagger - \epsilon < \hat{q}_{\rm Id}$, i.e., the second optimal objective value goes outsides its feasible interval. Define this k as K and the algorithm terminates. We want quantify K shortly.

Lemma 3: Q_k, $k = 1, 2, \cdots, K$ are Pareto-optimal points.

A proof of Lemma 3 is given in Appendix D. The following property gives a lower bound for the gap between $\hat{q}_{k-1}^\dagger$ and $\hat{q}_k^\dagger$ for $k = 1, 2, \cdots, K$.

Property 2: For any two consecutive Pareto-optimal points Q_{k-1} and Q_k, $k = 1, 2, \cdots, K$, we have $\hat{q}_{k-1}^\dagger - \hat{q}_k^\dagger \ge \epsilon$.

A proof of Property 2 is given in Appendix E. Since the minimum vertical gap in each iteration is at least ϵ and that the interval for the second optimal objective values between Q_0 and Q_Z is $(\hat{q}_{\rm Nd} - \hat{q}_{\rm Id})$, we have an upper bound for K, the maximum number of iterations.

Corollary 2: $K \le \left\lceil \frac{\hat{q}_{\rm Nd} - \hat{q}_{\rm Id}}{\epsilon} \right\rceil$.

Phase II: Finding a Second Subset of Pareto-optimal Points. In Phase I, we find a subset of Pareto-optimal points by iteratively reducing the second optimal objective values. In Phase II, we can find a second subset of Pareto-optimal points by iteratively reducing the first optimal objective values. The details are similar to those in the preceding paragraphs and are omitted here to conserve space. In conclusion, we can find $Q_{K+1}(q_{K+1}^\dagger, \hat{q}_{K+1}^\dagger)$, $Q_{K+2}(q_{K+2}^\dagger, \hat{q}_{K+2}^\dagger)$,

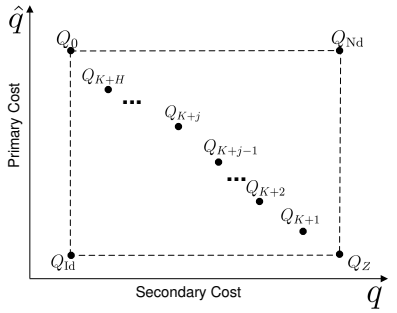


Fig. 4: The subset of Pareto-optimal points found in Phase II.

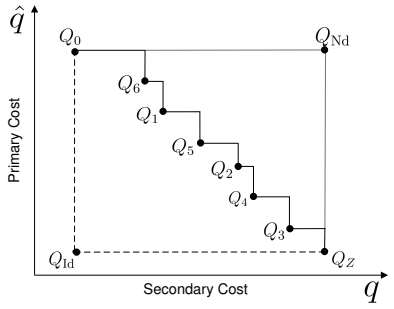


Fig. 5: An example to connect Pareto-optimal points in $\mathcal{H}$ to approximate the minimum cost curve.

$\cdots$, $Q_{K+H}(q^\dagger_{K+H}, \hat{q}^\dagger_{K+H})$ based on $Q_Z(q_{\rm Nd}, \hat{q}_{\rm Nd})$, $Q_{K+1}(q^\dagger_{K+1}, \hat{q}^\dagger_{K+1})$, $\cdots$, $Q_{K+H-1}(q^\dagger_{K+H-1}, \hat{q}^\dagger_{K+H-1})$, respectively, where H is the number of iterations until $q^\dagger_{K+H} - \epsilon < q_{\rm Id}$. Denote $\{Q_Z, Q_{K+1}, Q_{K+2}, \cdots, Q_{K+H}, Q_0\}$ as the Pareto-optimal point subset $\mathcal{H}_2$ (see Fig. 4). We can prove that Q_{K+j} are Pareto-optimal points with $j = 1, 2, \cdots, H$; $q^\dagger_Z - q^\dagger_{K+1} \geq \epsilon$ and $q^\dagger_{K+j-1} - q^\dagger_{K+j} \geq \epsilon$ for any two consecutive Pareto-optimal points Q_{K+j-1} and Q_{K+j} with $j = 2, 3, \cdots, H$; and the maximum number of iterations for H is $H \leq \lceil \frac{q_{\rm Nd} - q_{\rm Id}}{\epsilon} \rceil$.

Merging Two Subsets of Pareto-optimal points. Now we merge $\mathcal{H}_1$ and $\mathcal{H}_2$ into one subset of Pareto-optimal points and remove any duplication if any. Denote the newly merged subset as $\mathcal{H}$.

4.4 Connecting Pareto-optimal Points

Recall that the purpose of finding a subset of Pareto-optimal points $\mathcal{H}$ between Q_0 and Q_Z is to connect these points to approximate the minimum cost curve. In this section, we show how to do this.

As an example, suppose that $\mathcal{H}_1 = \{Q_0, Q_1, Q_2, Q_3, Q_Z\}$ in Phase I, and $\mathcal{H}_2 = \{Q_Z, Q_4, Q_5, Q_6, Q_0\}$ in Phase II. We merge $\mathcal{H}_1$ and $\mathcal{H}_2$ into $\mathcal{H} = \mathcal{H}_1 \bigcup \mathcal{H}_2 = \{Q_0, Q_6, Q_1, Q_5, Q_2, Q_4, Q_3, Q_Z\}$. To approximate the minimum cost curve, we connect the two neighbouring points in $\mathcal{H}$ through staircase linear segments (see Fig. 5). In general, we connect the two neighbouring points in $\mathcal{H}$ whose coordinates are closest to each other through staircase linear segments to construct an approximation curve.

Theorem 1: The approximation curve constructed by the set of Pareto-optimal points in $\mathcal{H}$ is an ϵ-approximation to the minimum cost curve for problem BIOPT$(\mu, \hat{\mu})$.

A proof of Theorem 1 is given in Appendix F.

The first application of the approximation curve is that it shows the entire landscape of minimum cost value for a single objective (between the two) or both objectives with approximation performance guarantee. That is, using this curve, for any given cost for the primary (secondary) network, we can immediately find the minimum cost that will incur in the secondary (primary) network with approximation performance guarantee.

5 Scaling ϵ-approximate Curve for Different Cost Parameters

In Section 3.5, we presented an interesting finding (Property 1) that shows how to find a new minimum cost curve when the cost parameters μ and $\hat{\mu}$ vary. For the ϵ-approximation curve, we would like to find a similar property. This is stated in the following sequel property.

AS-OPT: An Approximate Solution to S-OPT

1. **Input:** ϵ-approximate curve for problem BIOPT$(\mu, \hat{\mu})$;
2. **Procedure:**
3. Step 1: Scale each point on the ϵ-approximation
4. curve for BIOPT$(\mu, \hat{\mu})$ by factors of (λ_1, λ_2)
5. to obtain the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve
6. for BIOPT$(\lambda_1\mu, \lambda_2\hat{\mu})$;
7. Step 2: Check all the Pareto-optimal points
8. $(\lambda_1 q^\dagger, \lambda_2 \hat{q}^\dagger)$ on the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation
9. curve to find the point with $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$.
10. Denote this point as $Q_{\rm S-OPT}$ and its solution
11. as $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$;
12. **Output:** The value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ and solution ω.

Fig. 6: An algorithm to offer an approximate solution to single objective optimization problem based on the ϵ-approximation curve.

Property 3: For a Pareto-optimal point $(q^\dagger, \hat{q}^\dagger)$ with a solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ on the ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$, if the cost parameters μ and $\hat{\mu}$ are scaled by factors of α_1 and α_2, respectively (i.e., $\mu \to \alpha_1\mu$ and $\hat{\mu} \to \alpha_2\hat{\mu}$), then we have a Pareto-optimal point $(\alpha_1 q^\dagger, \alpha_2 \hat{q}^\dagger)$ to the new problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ while a solution to $(\alpha_1 q^\dagger, \alpha_2 \hat{q}^\dagger)$ is the same as $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$.

The above property can be easily proved by comparing BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ to BIOPT$(\mu, \hat{\mu})$ and noticing that the two objectives in BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ are the scaled version of these two objectives in BIOPT$(\mu, \hat{\mu})$ while the constraints are identical.

The ϵ-approximation curve contains both weakly Pareto-optimal points and other points (feasible but not weakly Pareto-optimal). If we scale each point on the ϵ-approximation curve by factors of α_1 and α_2 (i.e., α_1 for $q_{\rm S}$ and α_2 for $\hat{q}_{\rm P}$), then what kind of curve are we getting? To answer this question, we first give the following definition.

Definition 1: The curve that is obtained by scaling each point on an ϵ-approximation curve by factors of α_1 and α_2 (i.e., α_1 for $q_{\rm S}$ and α_2 for $\hat{q}_{\rm P}$) is called an $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve.

With this definition, we present the following main result.

Theorem 2: The $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve is an approximation to the minimum cost curve for problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ in that for any point on the vertical segment of the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve, the approximation error is less than $\alpha_1\epsilon$ in the $q_{\rm S}$ dimension (horizonal axis), and for any point on the horizontal segment of the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve, the approximation error is less than $\alpha_2\epsilon$ in the $\hat{q}_{\rm P}$ dimension (vertical axis).

A proof of Theorem 2 is given in Appendix G. Note that based on Definition 1, the ϵ-approximation curve guarantees an error bound of ϵ in both $q_{\rm S}$ and $\hat{q}_{\rm P}$ dimensions and can also be denoted as (ϵ, ϵ)-approximation curve.

6 Extension to Single Objective Optimization Problems

As another application of the ϵ-approximate curve obtained in the previous section, we now show how it can be extended to serve as an approximate solution to a single objective

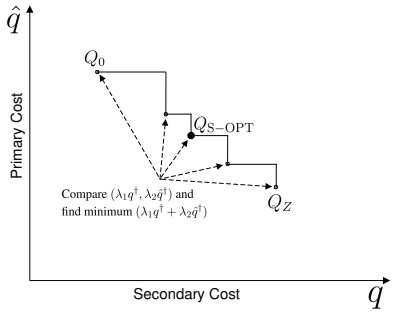


Fig. 7: Finding $Q_{\mathrm{S-OPT}}$ to offer an approximate solution to S-OPT.

optimization problem. To see how this works, let us consider the following weighted cost (single objective) optimization problem:

S-OPT

min $\lambda_1 q_{\mathrm{S}} + \lambda_2 \hat{q}_{\mathrm{P}}$

s.t. Mutual interference constraints: (4);
Self-interference constraints: (6);
Primary rate requirements: (7), (8);
Secondary rate requirements: (10), (11);
Link flow constraints: (13);

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ and the set of the constraints in S-OPT is identical to those in BIOPT$(\mu, \hat{\mu})$. Since $q_{\mathrm{S}} = \mu\hat{\varphi}$ and $\hat{q}_{\mathrm{P}} = \hat{\mu}\varphi$, we can rewrite the objective $\min(\lambda_1 q_{\mathrm{S}} + \lambda_2 \hat{q}_{\mathrm{P}})$ as $\min(\lambda_1 \mu\hat{\varphi} + \lambda_2 \hat{\mu}\varphi)$. Now, we give a simple algorithm that takes advantage of the ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$ to solve this single objective optimization problem.

Fig. 6 gives the pseudo-code of this simple algorithm, which we call AS-OPT. To show how it works, suppose that we have an ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$. In step 1 of the algorithm, we scale the cost parameters μ and $\hat{\mu}$ by factors of λ_1 and λ_2, respectively (i.e., $\mu \to \lambda_1\mu$ and $\hat{\mu} \to \lambda_2\hat{\mu}$) and obtain the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve for problem BIOPT$(\lambda_1\mu, \lambda_2\hat{\mu})$ per Definition 1. In step 2 of the algorithm, we check all the Pareto-optimal points $(\lambda_1 q^\dagger, \lambda_2 \hat{q}^\dagger)$ on the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve to find the point with $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ among all the Pareto-optimal points and denote it as $Q_{\mathrm{S-OPT}}$ (see Fig. 7). Finally, the algorithm gives an output $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ and $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$.

Theorem 3: A feasible solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ in the output of the AS-OPT algorithm is also a feasible solution to problem S-OPT. Further, the approximation error between the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ by AS-OPT and the optimal objective for the S-OPT is upper bounded by $(\lambda_1 + \lambda_2)\epsilon$.

A proof of Theorem 3 is given in Appendix H.

7 Numerical Results

In this section, we present numerical results to validate the ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$ and demonstrate its applications.

7.1 Network Setting

As a case study, we consider a randomly generated 15-node primary network and a 15-node secondary network in a 100×100 area (see Fig. 8). For generality, we normalize

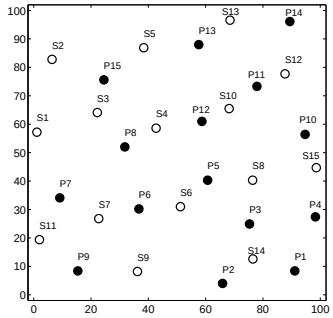


Fig. 8: A 15-node primary network and a 15-node secondary network.

TABLE 2: Source-destination nodes and rate requirements for each session in the network.

Sessions	Source	Destination	Rate Req.
Primary session 1	P_{10}	P_{14}	1.5
Primary session 2	P_7	P_3	1.8
Secondary session 1	S_{11}	S_3	1.8
Secondary session 2	S_1	S_6	1.5

all units for distance, data rate, bandwidth and power with appropriate dimensions. The location of each node is randomly generated. We assume the transmission range and interference range for each node are 30 and 50, respectively. We assume that the primary and secondary networks share the same pool of channels with a total bandwidth (W) 10 and that a time frame has eight time slots ($T = 8$). The power spectral density ρ_i for each node $i \in \mathcal{N}$ is 1, the path loss index γ is 4 and the ambient Gaussian noise density N_0 is 10^{-6}. The link's capacity between transmit node i and receive node j is calculated by $C_{ij} = W \log_2(1 + \frac{\rho_i d_{ij}^{-\gamma}}{N_0})$, where d_{ij} is the distance between nodes i and j.

We assume that there are two primary sessions in the primary network and two secondary sessions in the secondary network. The source and destination nodes for each session are randomly chosen and are shown in Table 2. Rate requirement for each primary and secondary sessions is also given in Table 2. We set $\mu = 1$ and $\hat{\mu} = 1$ for the cost parameters. We set the target approximation error $\epsilon = 0.05$.

7.2 Validation of ϵ-approximation Curve

Under the above network setting, we apply our algorithm to iteratively find a subset of Pareto-optimal points to construct an ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$.

We used a commercial solver (CPLEX) to solve our Mixed Integer Linear Programming (MILP). The CPLEX solver was run on a Dell Precision T7600 workstation, with dual Intel Xeon CPUE5-2687W CPUs (each with 8 cores) running at 3.1 GHz. The memory of the workstation is 64 GB and the OS is Windows 7 Professional. We find the initial Pareto-optimal point $Q_0(0, 3.3)$ and the final Pareto-optimal point $Q_Z(5.4, 0)$. In Phase I, we find $\mathcal{H}_1 = \{(0, 3.3), (0.038, 1.8), (1.8, 0.966), (3.198, 0.543), (3.248, 0.493), (3.298, 0.443), (5.4, 0)\}$, and in Phase II, we find $\mathcal{H}_2 = \{(5.4, 0), (3.343, 0.398), (3.293, 0.448), (3.243, 0.498), (1.8, 0.966), (0.038, 1.8), (0, 3.3)\}$. Then we have $\mathcal{H} = \mathcal{H}_1 \bigcup \mathcal{H}_2 = \{(0, 3.3), (0.038, 1.8), (1.8, 0.966), (3.198, 0.543), (3.243, 0.498), (3.248, 0.493), (3.293, 0.448),$

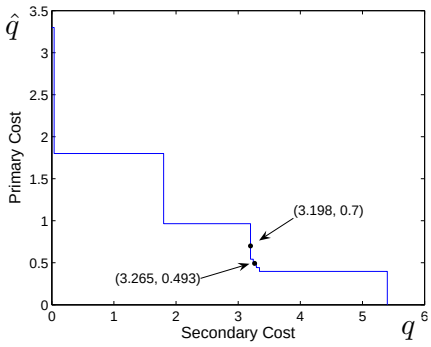


Fig. 9: An ϵ-approximation curve for problem BIOPT$(\mu, \hat{\mu})$ found by our algorithm.

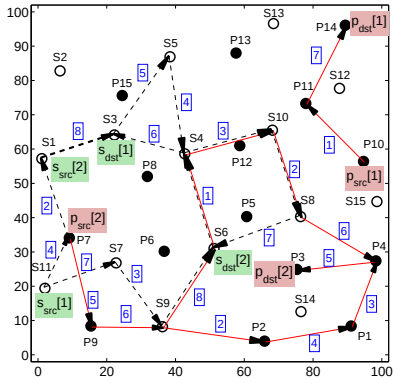


Fig. 10: Solution for flow routing and scheduling corresponding to weakly Pareto-optimal Point $(3.198, 0.7)$. A number in the box indicates the time slot when the link is active.

$(3.298, 0.443)$, $(3.343, 0.398)$, $(5.4, 0)\}$. Based on $\mathcal{H}$, we plot an ϵ-approximation curve in Fig. 9. The run-time for calculating one Pareto-optimal point in ϵ-approximation curve is about 2 hours for epsilon=0.05, and the total runtime for the entire ϵ-approximation curve is about 20 hours.

Now we validate that the constructed curve in Fig. 9 is indeed an ϵ-approximation curve (Theorem 1). We follow the following process. For any point $(q^*, \hat{q}^*)$ on a vertical segment of the constructed curve in Fig. 9, we can compute the first objective function $(\min q_{\text{S}})$ while assigning $\hat{q}_{\text{P}} = \hat{q}^*$ in BIOPT$(\mu, \hat{\mu})$. This is a single objective optimization problem and we can find $q^\dagger$. If $q^* - q^\dagger < \epsilon$, we can claim that this point $(q^*, \hat{q}^*)$ on the vertical segment is an ϵ-approximation to point $(q^\dagger, \hat{q}^*)$, which is on the minimum cost curve. Likewise, for any point $(q^*, \hat{q}^*)$ on a horizontal segment of the constructed curve in Fig. 9, we can compute the second objective function $(\min \hat{q}_{\text{P}})$ while assigning $q_{\text{S}} = q^*$ in BIOPT$(\mu, \hat{\mu})$. This is a single objective optimization problem and we can find $\hat{q}^\dagger$. If $\hat{q}^* - \hat{q}^\dagger < \epsilon$, we can claim that this point $(q^*, \hat{q}^*)$ on the horizontal segment is an ϵ-approximation to point $(q^*, \hat{q}^\dagger)$, which is on the minimum cost curve. Putting the two cases together, we can verify that any point on the constructed curve is an ϵ-approximation to a point on the minimum cost curve for problem BIOPT$(\mu, \hat{\mu})$.

We randomly pick two points for illustration, with the first point on the vertical segment and the second point on the horizontal segment of the constructed curve in Fig. 9. For the first point, we pick $(3.198, 0.7)$. We can compute the first objective function $(\min q_{\text{S}})$ while assigning $\hat{q}_{\text{P}} = \hat{q}^* = 0.7$ in BIOPT$(\mu, \hat{\mu})$ and find $q^{\dagger} \cdot = 3.198$. Then the approximation error is $q^* - q^\dagger = 0 < \epsilon$, which is correct. A solution for flow routing and scheduling corresponding to this point $(3.198, 0.7)$ is shown in Fig. 10. Secondary cost q^* occurs when secondary nodes S_4, S_6, S_8, S_9 and S_{10} relay the primary traffic (for primary session 2). According to the simulation results, $\hat{f}_{S_4,S_{10}}(2) = 0.349$, $\hat{f}_{S_6,S_4}(2) = 0.349$, $\hat{f}_{S_8,P_4}(2) = 0.349$, $\hat{f}_{S_9,P_2}(2) = 1.451$, $\hat{f}_{S_9,S_6}(2) = 0.349$, and $\hat{f}_{S_{10},S_8}(2) = 0.349$. Since $\mu = 1$ and $\hat{\varphi} = \hat{f}_{S_4,S_{10}}(2) + \hat{f}_{S_{10},S_8}(2) + \hat{f}_{S_8,P_4}(2) + \hat{f}_{S_9,S_6}(2) + \hat{f}_{S_9,P_2}(2) + \hat{f}_{S_6,S_4}(2) = 3.198$, we have $q^* = \mu\hat{\varphi} = 3.198$. Meanwhile, primary cost $\hat{q}^*$ occurs when primary node P_7 relays the secondary traffic (for secondary session 1). Since $\hat{\mu} = 1$ and $\varphi = f_{P_7,S_1}(1) = 0.7$, we have $\hat{q}^* = \hat{\mu}\varphi = 0.7$. Hence, this solution for flow routing and scheduling indeed matches point $(3.198, 0.7)$. For the second point, we pick $(3.265, 0.493)$. Following the same validation method, we can compute the second objective function $(\min \hat{q}_{\text{P}})$ while assigning $q_{\text{S}} = q^* = 3.265$ in BIOPT$(\mu, \hat{\mu})$ and find $\hat{q}^\dagger = 0.475$. The approximation error is $\hat{q}^* - \hat{q}^\dagger = 0.018 < \epsilon$, which is correct.

We can repeat the validation for any other points on the vertical and horizontal segments on the curve in Fig. 9 and obtain the same conclusion. Therefore, the constructed curve is an ϵ-approximation to the minimum cost curve for problem BIOPT$(\mu, \hat{\mu})$.

For the case of varying target approximation error ϵ, the results can be obtained by executing our algorithm. New simulation results with target approximation error $\epsilon = 0.1$ are shown in Fig. 11, where the ϵ-approximation curves for $\epsilon = 0.1$ and $\epsilon = 0.05$ are compared. These two stair curves are similar and almost overlapping as shown in Fig. 11, but there are more Pareto-optimal points when $\epsilon = 0.05$.

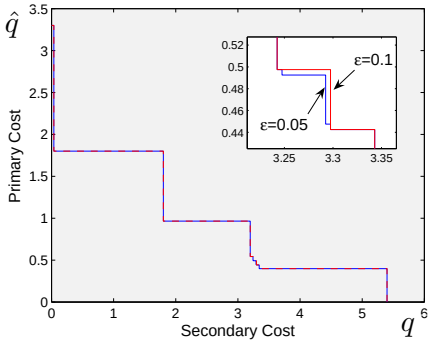
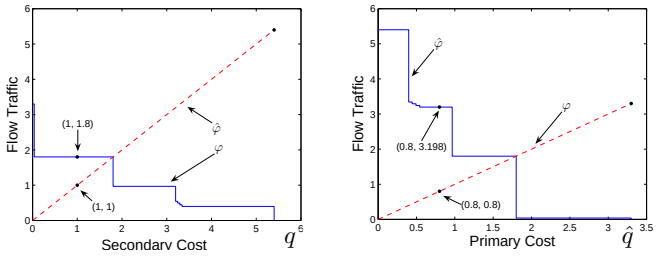


Fig. 11: A comparison of the ϵ-approximation curve between $\epsilon = 0.1$ and $\epsilon = 0.05$

7.3 Applications of ϵ-approximation curve

Finding minimum relayed traffic for a given cost. Based on the ϵ-approximation curve in Fig. 9 and the linear relationship between secondary (primary) cost and primary (secondary) traffic relayed by the secondary (primary) nodes defined in BIOPT, i.e., $q_{\text{S}} = \mu \cdot \hat{\varphi}$ $(\hat{q}_{\text{P}} = \hat{\mu} \cdot \varphi)$, we can immediately plot Fig. 12. The staircase curve in Fig. 12(a) shows the minimum amount of secondary traffic φ that can be relayed by the primary network for a given cost on the secondary network; the dashed line in Fig. 12(a) shows the amount of primary traffic $\hat{\varphi}$ that can be relayed by the secondary network for a



(a) (b)

Fig. 12: (a) Finding minimum amount of secondary traffic that can be relayed by primary nodes for a given q on the staircase curve and the amount of primary traffic that can be relayed by secondary nodes for a given q on the dashed line. (b) Finding minimum amount of primary traffic that can be relayed by secondary nodes for a given $\hat{q}$ on the staircase curve and the amount of secondary traffic that can be relayed by primary nodes for a given $\hat{q}$ on the dashed line.

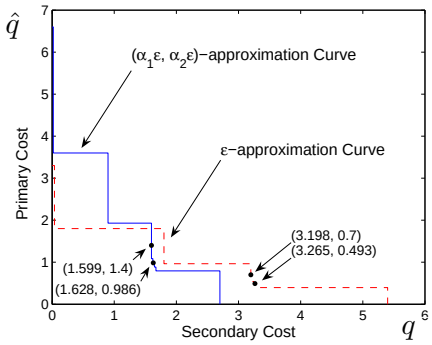
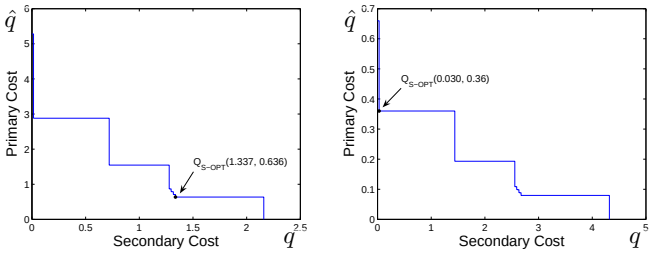


Fig. 13: Scaling for an $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve from an ϵ-approximation one.

given cost on the secondary network. Note that the staircase curve for φ is obtained by $\hat{q}$ in Fig. 9 while the dashed line for $\hat{\varphi}$ is obtained by q. Fig. 12(b) is obtained in the same fashion as Fig. 12(a) with the only difference being that the horizontal axis is now $\hat{q}$, the counterpart of q in Fig. 9. The results in Fig. 12 (a) and (b) are very useful as one can read out the minimum amount of secondary (primary) traffic that can be relayed by the primary (secondary) nodes for a given q ($\hat{q}$) on the staircase curve and the amount of primary (secondary) traffic that can be relayed by the secondary (primary) nodes for a given q ($\hat{q}$) on the dashed line. For example, in Fig. 12(a), for $q = 1$, we have that a minimum of 1.8 unit of secondary traffic can be relayed by the primary nodes on the staircase curve while 1 unit of primary traffic can be relayed by the secondary nodes on the dashed line. In Fig. 12(b), for $\hat{q} = 0.8$, we have that a minimum of 3.198 unit of primary traffic can be relayed by the secondary nodes on the staircase curve while 0.8 unit of secondary traffic can be relayed by the primary nodes on the dashed line.

Validation of $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation Curve. Now we validate Theorem 2 numerically. Suppose the scaling factors $\alpha_1 = 0.5$ and $\alpha_2 = 2$. Then based on Definition 1, for the ϵ-approximation curve in Fig. 9, we can obtain an $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13 (solid staircase line segments). The original ϵ-approximation curve in Fig. 9 is also shown in Fig. 13 for scaling comparison.

Now we validate that the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13 is indeed an approximation to the minimum cost curve for problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ (Theorem 2). Similar



(a) (b)

Fig. 14: (a) Finding $Q_{\mathrm{S-OPT}}(1.337, 0.636)$ to offer an approximate solution to S-OPT $\min(0.4q_{\mathrm{S}} + 1.6\hat{q}_{\mathrm{P}})$. (b) Finding $Q_{\mathrm{S-OPT}}(0.030, 0.36)$ to offer an approximate solution to S-OPT $\min(0.8q_{\mathrm{S}} + 0.2\hat{q}_{\mathrm{P}})$.

to the validation method of the ϵ-approximation curve in Section 7.3, for any point $(\alpha_1 q^*, \alpha_2\hat{q}^*)$ on a vertical segment of the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13, we can compute the first objective function $(\min \alpha_1 q_{\mathrm{S}})$ while assigning $\alpha_2\hat{q}_{\mathrm{P}} = \alpha_2\hat{q}^*$ in BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$. This is a single objective optimization problem and we can find $\alpha_1 q^\dagger$. If $\alpha_1 q^* - \alpha_1 q^\dagger < \alpha_1\epsilon$, we can claim that this point $(\alpha_1 q^*, \alpha_2\hat{q}^*)$ on the vertical segment is an $\alpha_1\epsilon$-approximation to point $(\alpha_1 q^\dagger, \alpha_2\hat{q}^*)$, which is on the minimum cost curve. Likewise, for any point $(\alpha_1 q^*, \alpha_2\hat{q}^*)$ on a horizontal segment of the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13, we can compute the second objective function $(\min \alpha_2\hat{q}_{\mathrm{P}})$ while assigning $\alpha_1 q_{\mathrm{S}} = \alpha_1 q^*$ in BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$. This is also a single objective optimization problem and we can find $\alpha_2\hat{q}^\dagger$. If $\alpha_2\hat{q}^* - \alpha_2\hat{q}^\dagger < \alpha_2\epsilon$, we can claim that this point $(\alpha_1 q^*, \alpha_2\hat{q}^*)$ on the horizontal segment is an $\alpha_2\epsilon$-approximation to point $(\alpha_1 q^*, \alpha_2\hat{q}^\dagger)$, which is on the minimum cost curve. Putting the two cases together, we can verify that any point on the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve is an approximation to a point on the minimum cost curve for problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$.

We randomly pick two points for illustrations, with the first point on the vertical segment and the second point on the horizontal segment of the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13. For the first point, we pick $(1.599, 1.4)$, which is obtained by scaling the point $(3.198, 0.7)$ on the ϵ-approximation for problem BIOPT$(\mu, \hat{\mu})$ by (0.5, 2). We can compute the first objective function $(\min \alpha_1 q_{\mathrm{S}})$ while assigning $\alpha_2\hat{q}_{\mathrm{P}} = \alpha_2\hat{q}^* = 1.4$ in BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ and find $\alpha_1 q^\dagger = 1.599$. Then the approximation error is $\alpha_1 q^* - \alpha_1 q^\dagger = \alpha_1 \times 0 = 0 < \alpha_1\epsilon$, which is correct. For the second point, we pick $(1.628, 0.986)$, which is obtained by scaling the point $(3.265, 0.493)$ on the ϵ-approximation for problem BIOPT$(\mu, \hat{\mu})$ by (0.5, 2). Following the same validation method, we can compute the second objective function $(\min \alpha_2\hat{q}_{\mathrm{P}})$ while assigning $\alpha_1 q_{\mathrm{S}} = \alpha_1 q^* = 1.628$ in BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$ and find $\alpha_2\hat{q}^\dagger = 0.95$. The approximation error is $\alpha_2\hat{q}^* - \alpha_2\hat{q}^\dagger = \alpha_2 \times 0.018 = 0.036 < \alpha_2\epsilon$, which is correct.

We can repeat the validation for any other points on the vertical and horizontal segments on the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation curve in Fig. 13 and obtain the same conclusion. Therefore, the $(\alpha_1\epsilon, \alpha_2\epsilon)$-approximation is an approximation to the minimum cost curve for problem BIOPT$(\alpha_1\mu, \alpha_2\hat{\mu})$.

Extension to Single Objective Optimization Problems. In the following, we validate Theorem 3. We follow the following process. First, we validate a feasible solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$

in the output of the AS-OPT algorithm (see Fig. 6) is a feasible solution to problem S-OPT $\min(\lambda_1 q_{\rm S} + \lambda_2 \hat{q}_{\rm P})$. As the output of the AS-OPT algorithm, the solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ corresponding to a Pareto-optimal point $Q_{\rm S-OPT}$ on the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve is feasible after we obtained the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve for problem BIOPT$(\lambda_1\mu, \lambda_2\hat{\mu})$ (by scaling each point on the ϵ-approximate curve for problem BIOPT$(\mu, \hat{\mu})$ by factors of λ_1 and λ_2). Since problem BIOPT$(\lambda_1\mu, \lambda_2\hat{\mu})$ and problem S-OPT have the same set of constraints in their formulations, the feasible solution $\omega = \{\mathbf{x}, \mathbf{f}, \hat{\mathbf{f}}\}$ in the output of the AS-OPT algorithm is also a feasible solution to problem S-OPT.

We now validate that the approximation error between the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ by AS-OPT algorithm and the optimal objective for problem S-OPT is upper bounded by $(\lambda_1 + \lambda_2)\epsilon$. In AS-OPT algorithm, after we obtained the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve for problem BIOPT$(\lambda_1\mu, \lambda_2\hat{\mu})$ (by scaling each point on the ϵ-approximate curve for problem BIOPT$(\mu, \hat{\mu})$ by factors of λ_1 and λ_2), we check all the Pareto-optimal points $(\lambda_1 q^\dagger, \lambda_2 \hat{q}^\dagger)$ on the $(\lambda_1\epsilon, \lambda_2\epsilon)$-approximation curve to find the point $Q_{\rm S-OPT}$ with $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$. Hence, we have the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ in AS-OPT algorithm. For comparison, we can solve problem S-OPT and find its optimal objective value. If the difference between the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger)$ in AS-OPT algorithm and the optimal objective value for problem S-OPT is less than $(\lambda_1 + \lambda_2)\epsilon$, we are done.

For illustrations, we randomly pick two sets of λ_1 and λ_2. For the first set, we have $\lambda_1 = 0.4$ and $\lambda_2 = 1.6$. Following our validation method, in AS-OPT algorithm, we find the Pareto-optimal point $Q_{\rm S-OPT}(1.337, 0.636)$ (see Fig. 14(a)) with $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger) = 1.973$. So the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger) = 1.973$ in AS-OPT algorithm. On the other hand, we solve problem S-OPT and find its optimal objective value is also 1.973. Then the approximation error between the two is 0 and is less than $(\lambda_1 + \lambda_2)\epsilon = 0.1$, which is correct.

For the second set, we have $\lambda_1 = 0.8$ and $\lambda_2 = 0.2$. Following the same validation method, in AS-OPT algorithm, we find the Pareto-optimal point $Q_{\rm S-OPT}(0.030, 0.36)$ (see Fig. 14(b)) with $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger) = 0.390$. So the output value $\min(\lambda_1 q^\dagger + \lambda_2 \hat{q}^\dagger) = 0.390$ in AS-OPT algorithm. On the other hand, we solve problem S-OPT and find its optimal objective value is also 0.390. Then the approximation error between the two is 0 and is less than $(\lambda_1 + \lambda_2)\epsilon = 0.05$, which is correct.

For any other sets of λ_1 and λ_2, we can validate and obtain the same conclusion.

7.4 Other Case Studies

In addition to the case study in Section 7.3 with the network setting in Section 7.1, we also vary network topologies, number of nodes for primary and secondary networks, number of network sessions and the rate requirements, the simulation results are all consistent and we have the same conclusions.

8 Conclusions

In this paper, we offered an in-depth study on potential costs incurred in the primary and secondary networks when they are allowed to relay each other's traffic. We presented cost models for cooperative traffic relaying between the two networks and studied a multiobjective optimization problem. The main contributions of this paper are: (i) a two-phase iterative algorithm to find a subset of Pareto-optimal points that can be used to construct an ϵ-approximation curve; (ii) an application of the ϵ-approximation curve to show the entire landscape of minimum cost value for a single objective (between the two) or both objectives with approximation performance guarantee; (iii) an application of the ϵ-approximation curve to find a new approximate minimum cost curve for a different set of cost parameters; (iv) an application of the ϵ-approximation curve to find an approximate solution for a single objective optimization problem. Collectively, these results offer important theoretical foundation on quantifying mutual cost in the primary and secondary networks when they help relay each other's traffic in a cooperative manner.

References

[1] J.F. Berube, M. Gendreau, and J.Y. Potvin, "An exact ϵ-constraint method for bi-objective combinatorial optimization problems: Application to the Traveling Salesman Problem with Profits," *European Journal of Operational Research,* vol. 194, no. 1, pp. 39–50, Apr. 2009.

[2] N.D. Chatzidiamantis, E. Matskani, L. Georgiadis, I. Koutsopoulos and L. Tassiulas, "Optimal Primary-Secondary User Cooperation Policies in Cognitive Radio Networks," *IEEE Trans. on Wireless Commun.,* vol. 14, no. 6, pp. 3443–3455, Jun. 2015.

[3] D. Chiarotto, O. Simeone, and M. Zorzi, "Spectrum Leasing via Cooperative Opportunistic Routing Techniques," *IEEE Trans. on Wireless Commun.,* vol. 10, no. 9, pp. 2960–2970, Sept. 2011.

[4] J.L. Cohon, *Multiobjective Programming and Planning,* Courier Dover, New York, 2004.

[5] L. Duan, L. Gao and J. Huang, "Cooperative Spectrum Sharing: A Contract-based Approach," *IEEE Trans. on Mobile Computing,* vol. 13, no. 1, pp. 174–187, Jan. 2014.

[6] M. Ehrgott, *Multicriterion Optimization,* Springer-Verlag, New York, 2005.

[7] I. Filippini, and M. Cesana, "A New Outlook on Routing in Cognitive Radio Networks: Minimum-Maintenance-Cost Routing," *IEEE Trans. on Networking,* vol. 21, no. 5, pp. 1484–1498, Oct. 2013.

[8] S. Gass, and T. Saaty, "The computational algorithm for the parametric objective function," *Naval Research Logistics Quarterly,* vol. 2, no. 1, pp. 39–45, Mar. 1995.

[9] S. Geirhofer, L. Tong, and B.M. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Communications Magazine,* vol. 45, no. 5, pp. 66–72, May 2007.

[10] A. Goldsmith, S.A. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock with Cognitive Radios: An Information Theoretic Perspective," *Proceedings of the IEEE,* vol. 97, no. 5, pp. 894–914, May 2009.

[11] Y.T. Hou, Y. Shi, and R. Sherali, "Spectrum Sharing for Multi-Hop Networking with Cognitive Radios," *IEEE Journal on Selected Areas in Commun.,* vol. 26, no. 1, pp. 146–155, Jan. 2008.

[12] Y.T. Hou, Y. Shi, and R. Sherali, *Applied Optimization Methods For Wireless Networks.* Chapter 5, Cambridge University Press, 2014.

[13] S. Hua, H. Liu, M. Wu, X. Zhuo, M. Wu and S.S. Panwar, "Exploiting Multiple Antennas in Cooperative Cognitive Radio Networks," *IEEE Trans. on Vehicular Technology,* vol. 63, no. 7, pp. 3318–3330, Sept. 2014.

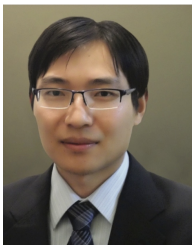
[14] S.K. Jayaweera, M. Bkassiny, and K. Avery, "Asymmetric cooperative communication based spectrum leasing via auctions in cognitive radio networks," *IEEE Trans. on Wireless Commun.,* vol. 10, no. 8, pp. 2716–2724, Aug. 2011.

[15] M. Karaca, K. Khalil, E. Ekici, and O. Ercetin, "Optimal Scheduling and Power Allocation in Cooperate-to-Join Cognitive Radio Networks," *IEEE/ACM Trans. on Networking,* vol. 21, no. 6, pp. 1708–1721, Dec. 2013.

[16] I.Y. Kim, and O.L. Weck, "Adaptive weighted sum method for bi-objective optimization: Pareto front generation," *Structural and Multidisciplinary Optimization,* vol. 29, no. 9, pp. 149–158, Jan. 2005.

[17] G. Kirlik, and S. Sayin, "A new algorithm for generating all non-dominated solutions of multiobjective discrete optimization problems," *European Journal of Operational Research,* vol. 232, no. 3, pp. 479–488, Feb. 2014.
[18] W. Li, X. Cheng, T. Jing and X. Xing, "Cooperative Multi-hop Relaying via Network Formation Games in Cognitive Radio Networks," in *Proc. IEEE INFOCOM,* pp. 971–979, Turin, Italy, Apr. 14-19, 2013.
[19] Q. Liang, X. Wang, X. Tian, and Q. Zhang, "Route-Switching Games in Cognitive Radio Networks," in *Proc. ACM MobiHoc,* pp. 249–252, Bangalore, India, Jul. 29-Aug. 1, 2011.
[20] R. Manna, R.H.Y. Louie, Y. Li, and B. Vucetic, "Cooperative spectrum sharing in cognitive radio networks with multiple antennas," *IEEE Trans. on Signal Processing,* vol. 59, no. 11, pp. 5509–5522, Nov. 2011.
[21] T. Nadkar, V. Thumar, G. Shenoy, A. mehta, U.B. Desai, and S.N. Mechant, "A cross-layer framework for symbiotic relaying in cognitive radio networks," in *Proc. IEEE DySPAN,* pp. 498–509, Aachen, Germany, May 3-6, 2011.
[22] T.K. Ralphs, M.J. Saltzman, and M.M. Wiecek, "An Improved Algorithm for Biobjective Integer Programming," *Annals of Operations Research,* vol. 147, pp. 43–70, 2006.
[23] S. Ruzika, and M.M. Wiecek, "Approximation Methods in Multiobjective Programming," *Journal of Optimization Theory and Applications,* vol. 126, no. 3, pp. 473–501, Sept. 2005.
[24] W. Saad, Z. Han, H.V. Poor, T. Basar, and J.B. Song, "Cooperative Bayesian Nonparametric Framework for Primary User Activity Monitoring in Cognitive Radio Networks," *IEEE Journal on Selected Areas in Commun.,* vol. 30, no. 9, pp. 1815–1822, Oct. 2012.
[25] O. Simone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE Journal on Selected Areas in Commun.,* vol. 26, no. 1, pp. 203–213, Jan. 2008.
[26] W. Su, J.D. Natyjas, and S. Batalama, "Active Cooperation Between Primary Users and Cognitive Radio Users in Heterogeneous Ad-Hoc Networks' *IEEE Trans. on Signal Processing,* vol. 60, no. 4, pp. 1796–1805, Apr. 2012.
[27] L. Wang, and V. Fodor, "Dynamic Cooperative Secondary Access in Hierarchical Spectrum Sharing Networks," *IEEE Trans. on Wireless Commun.,* vol. 13, no. 11, pp. 6068–6080, Nov. 2014.
[28] A.M. Wyglinski, M. Nekovee, and Y.T. Hou, *Cognitive Radio Communications and Networks: Principles and Practice.* Academic Press/Elsevier, 2010.
[29] H. Xu, and B. li, "Resource Allocation with Flexible Channel Cooperation in Cognitive Radio Networks," *IEEE Trans. on Mobile Computing,* vol. 12, no. 5, pp. 957–970, May 2013.
[30] Y. Yan, J. Huang, and J. Wang, "Dynamic Bargaining for Relay-Based Cooperative Spectrum Sharing," *IEEE Journal on Selected Areas in Commun.,* vol. 31, no. 8, pp. 1480–1493, Aug. 2013.
[31] X. Yuan, Y. Shi, X. Qin, Y.T. Hou, W. Lou, S. Kompella, S.F. Midkiff, and J. H. Reed, "Beyond overlay: Reaping mutual benefits for primary and secondary networks through node-level cooperation," *IEEE Transactions on Mobile Computing,* vol. 16, no. 1, pp. 2–15, Jan. 2017.
[32] X. Yuan, F. Tian, Y.T. Hou, W. Lou, H. Sherali, S. Kompella, and J. H. Reed, "Optimal Throughput Curve for Primary and Secondary Users with Node-level Cooperation," in *Proc. IEEE DySPAN,* pp. 1–9, Stockholm, Sweden, Sept. 29-Oct. 2, 2015.
[33] J. Zhang, and Q. Zhang, "Stackelberg game for utility-based cooperative radio network," in *Proc. ACM MobiHoc,* pp. 23–32, New Orleans, Louisiana, USA, May 2009.
[34] Q. Zhao, and B.M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine,* vol. 24, no. 3, pp. 79–89, May 2007.

Feng Tian (M'13) received the Ph.D. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2008. From 2013 to 2015, he was a Visiting Scholar at Virginia Polytechnic Institute and State University, Blacksburg, VA, USA. He is currently an Associate Professor at Nanjing University of Posts and Telecommunications, Nanjing, China. His research interest focuses on performance optimization and algorithm design for wireless networks, cognitive radio networks and big data driven optimization for wireless networks.



Xu Yuan (S'13–M'16) received the B.S. degree from Nankai University, China in 2009, and the Ph.D. degree from Virginia Tech, Blacksburg, VA, USA, in 2016. From 2016 to 2017, he was a Post-Doctoral Fellow at the University of Toronto, Canada. He is currently an Assistant Professor at the University of Louisiana at Lafayette, LA, USA. His research interest focuses on cloud computing security, algorithm design and optimization for spectrum sharing, coexistence, and cognitive radio networks.

Y. Thomas Hou (F'14) is Bradley Distinguished Professor of Electrical and Computer Engineering at Virginia Tech, Blacksburg, VA, USA. He received his Ph.D. degree in Electrical Engineering from New York University (NYU) Tandon School of Engineering in 1998. Prof. Hous research focuses on developing innovative solutions to complex problems that arise in wireless networks. He is an IEEE Fellow and an ACM Distinguished Scientist. He is Chair of IEEE INFOCOM Steering Committee and a Distinguished Lecturer of the IEEE Communications Society.

Wenjing Lou (F'15) is a Professor in the Department of Computer Science at Virginia Tech, Falls Church, VA, USA. She received her Ph.D. degree in Electrical and Computer Engineering from the University of Florida in 2003. She recently completed her Program Director IPA assignment at the US National Science Foundation. Her research interests include cyber security and wireless networks. She is on the editorial boards of a number of IEEE transactions. Prof. Lou is an IEEE Fellow and Steering Committee Chair of IEEE Conference on Communications and Network Security (CNS).

Zhen Yang received the B.Eng. and M.E. degrees from the Nanjing University of Posts and Telecommunications, China, in 1983 and 1988, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, China, in 1999. From 1992 to 1993, he was a Visiting Scholar at Bremen University, Germany. In 2003, he was an Exchange Scholar at the University of Maryland, College Park, USA. He is currently a Professor at Nanjing University of Posts and Telecommunications, China. His research interests include signal processing, communication systems and networks. He served as the Vice Chair of the Chinese Institute of Communications and was a member of the Editorial Board of several journals, including the Chinese Journal of Electronics, Journal on Communications, and China Communications.