Locating network monitors: complexity, heuristics, and coverage

Kyoungwon Suh *, Yang Guo†, Jim Kurose, Don Towsley,

Department of Computer Science, University of Massachusetts, 140 Governors Dr., Amherst, MA, U.S.A. 01003

† *Thomson Corporate Research, Princeton Lab, 2 Independence way, Princeton, NJ, U.S.A. 08540*

Abstract

There is an increasing interest in passive monitoring of IP flows at multiple locations within an IP network. The objective of such a distributed monitoring system is to sample packets belonging to a large fraction of IP flows in a cost-effective manner by carefully placing monitors and controlling their sampling rates. In this paper, we consider the problem of where to place monitors within the network and how to control their sampling. To address the tradeoff between monitoring cost and monitoring coverage, we propose and study minimum cost and maximum coverage problems under various budget constraints and in the presence of routing changes caused by link failures. We show that all of the problem formulations are NP-hard. We propose greedy heuristics, and show that the heuristics provide solutions quite close to the optimal solutions through experiments using synthetic and real network topologies. In addition, our experiments show that a small number of monitors often suffices to monitor most of the traffic in an entire IP network.

Key words: Network monitoring,, Mathematical programming / optimization, Approximation algorithms, Set Cover Problem, Maximum Coverage Problem, Minimum Cost Problem, Packet sampling

1 Introduction

Traffic measurement and monitoring are important in order to understand the performance of a network infrastructure and to efficiently manage network resources. In particular, a passive monitoring system can be used to study packet-level traffic, estimate packet-size distributions, estimate the fined-grained volume of network traffic with different attributes for usage-based charging, and more [10]. In practice, a monitor is placed inside a router or deployed as a standalone measurement box tapped into a communication link. Once a monitor is placed on a link, it may capture packets (either all packets or sampled packets) carried by the link depending on its specific sampling configuration. In order to observe a large fraction of a network's traffic, it is necessary to monitor multiple links concurrently since only a relatively small fraction of the traffic can be seen at any single measurement point in a large IP network. Placing a monitor on a link incurs a deployment cost that includes fixed cost components such as the monitor's hardware/software cost, a space cost, and a maintenance cost. Therefore, it is important that the number of placed monitors be kept as small as possible. In addition, the actual monitoring operation performed by a monitor factors into its operating cost. The per-packet operating cost of each monitor depends mainly

* Corresponding author.
Email addresses: kwsuh@cs.umass.edu (Kyoungwon Suh), Yang.Guo@thomson.net (Yang Guo), kurose@cs.umass.edu (Jim Kurose), towsley@cs.umass.edu (Don Towsley).

on the link speed. In the case that packet traces are transferred to a central location for further analysis, the per-packet operating cost of each monitor may also depend on the distance and congestion level of the network path between the monitor and the central location. Monitors with lower per-packet operating costs are thus preferred whenever possible.

In this paper, we consider the problem of sampling packets in a cost-effective manner by carefully placing monitors and controlling their sampling rates. We consider IP networks in which each IP flow is routed along a single path. Because of single path routing, we are able (to first order) to observe all packets in an IP flow by monitoring any one of the links on the flow's path. Roughly speaking, we have two conflicting optimization objectives. One objective is to maximize the fraction of IP flows being sampled and the other objective is to minimize the total monitoring cost. We begin by introducing several monitoring cost and coverage models for a distributed monitoring system. Based on these models, we then formulate several optimization problems for determining the deployment and sampling strategies of monitors. The solution of each of these problems determines the minimal number of monitors and their optimal locations under various constraints; the operating strategy determines the optimal flow sampling rate of each monitor. We show that all of the problems are NP-hard. Therefore, we propose greedy heuristics and show that they provide close to optimal solutions for several problems based on synthetic and real network topologies. As a second contribution, we determine the relationship between the number of monitors and the maximum coverage of flows. Using both synthetic topologies and real topologies, we show that a relatively small number of monitors are sufficient to sample a large fraction of all IP flows in the network.

Several recent efforts have addressed the monitor placement problem in IP networks. The use of active probing has been proposed to obtain Internet topology and performance measurements (such as link delay and the existence of faults). The location of these active measurement devices or beacons has been determined using various heuristics [18, 4, 20, 6]. Also, Horton et al. [15] determined the minimal number of required beacons in a network and their optimal locations in order to obtain an accurate topology map. While these works investigated active monitoring, the IPMon project at Sprint [13] deployed multiple passive packet-level monitors inside their network to capture IP headers. However, they consider neither the problem of monitor location nor the choice of sampling methodology in their monitoring architecture. Recently, in parallel with our work, M. Sharma et al. [23] proposed a heuristic for locating passive monitors. However, they consider neither operating costs nor a sampling mode of operation, and do not analyze the complexity of the formulated monitoring problems, as we do in this paper.

The remainder of this paper is structured as follows. In the next section, we define a graph-based model of the monitoring problem, and the performance metrics which include deployment cost, operating cost, and monitoring reward. Sections 3 and 4 formulate various budget-constrained and coverage-constrained monitoring problems for both sampling and non-sampling modes of monitoring along with their complexity analysis and greedy heuristics. In Section 5, we evaluate the proposed greedy heuristics and examine the coverage issue. In Section 6, we extend our proposed monitoring problems by considering routing changes caused by single link failures and also by considering stochastic flow arrival/departure processes. A summary of related work is presented in Section 7. We conclude with a summary of our results and a discussion of future work.

2 Problem setting

In this section, we propose three models for the distributed monitoring problem — a deployment cost model, an operating cost model, and a monitoring reward model. The interactions among these three models are explored in Sections 3 and 4, where we show that it is possible to achieve near-optimal monitor placement and sampling strategies under various budget constraints and monitoring objectives.

We represent an IP network as an undirected graph, $G(V, E)$, where V and $E \subseteq V \times V$ denote the set of nodes and links, respectively. We define a *traffic flow* to be a collection of packets that originate and terminate at the same nodes, sharing the same route in the graph, i.e., ingress router to egress router flow. Flows can also be defined at different granularities. For instance, a flow may represent a network-to-network flow [1], ingress router-to-egress router flow, host-to-

[1] PoP-to-PoP flow, AS-to-AS flow, and Customer-

Table 1
Parameters in the model.

Parameter	Definition
L	Set of feasible links where monitors can be deployed
D	Set of all flows in the network
S_i	Set of all flows carried by link i ($S_i \subseteq D$)
S	Set of all S_i's, $i \in L$ ($S_i \in S$)
y_i	$\{0,1\}$ indicator if a flow monitoring station is deployed at link i, $i \in L$.
f_i	Deployment cost of a monitor at link i
ρ_j	traffic demand of flow j (unit: packets/time or bytes/time)
c_i	Unit operating cost of monitor at link i (unit: cost/packet or cost/byte)
x_{ij}	$\{0,1\}$ indicator if flow j is being monitored by a monitor at link i, $i \in L$, $j \in D$.
x_j	$\{0,1\}$ indicator if flow j is being monitored, $j \in D$.
m_{ij}	Fraction of flow j sampled by a monitor at link i
m_j	Fraction of flow j sampled by all monitors
$u_j(\{m_{ij}\})$	Nondecreasing concave function of the fraction of flow j sampled by all monitors (in sampling mode)
$v_j(x_j)$	Function of whether flow j is fully captured by any monitor (in non-sampling mode)

host flow, or application-level flow. We have chosen the definition of flow as ingress router-to-egress router flow for the following analysis since we believe that it is the most interesting for traffic engineering. The conclusions drawn from our analysis apply equally well to other flow definitions. We summarize the important flow and monitoring parameters in Table 1.

The solution to a distributed monitoring problem consists of two parts, *(i)* a set of links $L, L \subset E$ at which to place a monitor and *(ii)* a monitoring strategy (e.g., sampling rate) at each monitor. In determining the number of monitors and sampling rate at each monitor, we are interested in the tradeoff between monitoring cost and monitoring coverage. Therefore, we first introduce general cost and reward models for a distributed monitoring system. We define the monitor deployment and operating costs, and the reward from flow monitoring as follows (using the notation from Table 1). Here, D and S_i denote a set of all flows in the network and a set of all flows carried by link i, respectively. Let y_i indicate whether a monitor is deployed at link i:

- **Deployment cost.** The deployment cost captures the cost associated with deploying a monitoring station. Let f_i denote the deployment cost of a monitor at link i. Since the cost to place a monitor at a specific link may depend on geographical location or accessibility, the deployment cost to monitor link i, f_i, can differ from link to link. Hence, the total deployment cost is:

$$C_D = \textstyle\sum_{i \in L} f_i y_i \qquad (1)$$

- **Operating cost.** The operating cost captures the cost associated with the monitor's sampling operation, under the assumption that a passive monitoring station is able to monitor traffic that traverses a link in both directions. Let c_i denote the unit operating cost of a monitor on link i. This could represent the cost of sampling a single packet at link i and the cost of transferring the sampled packet to a post-processing location in the network for further analysis. The value of c_i can depend on the monitor, e.g., because of link speed and the hop count of the path from the monitor to the post-processing location. Let m_{ij} and ρ_j denote the fraction of flow j sampled by a monitor at link i and the traffic demand of flow j, respectively. The total operating cost at link i is a function of the total amount of traffic from all flows, j, sampled at link i:

$$C_O = \textstyle\sum_{i \in L} y_i c_i \sum_{j \in D} \rho_j m_{ij} \qquad (2)$$

- **Monitoring reward.** The monitoring reward captures the benefit of traffic monitoring. Let utility function $v_j(x_j)$ denote the benefit gained from monitoring flow j, assuming that the monitor captures all packets. Similarly, let utility function

network-to-Customer-network flow may be typical examples of network-to-network flow.

$u_j(\{m_{ij}\})$ denote the benefit gained from monitoring flow j, assuming sampling mode operation of monitors. For example, $v_j(x_j)$ and $u_j(\{m_{ij}\})$ may simply represent the monitored traffic demand of flow j.

$$C_M = \sum_{j\in D} v_j(x_j) \quad (3)$$
$$C_M = \sum_{j\in D} u_j(\{m_{ij}\}) \quad (4)$$

We assume that no additional benefit can be gained by repeatedly monitoring the same traffic. Thus, $u_j(\{m_{ij}\})$ can be expressed as follows:

$$u_j = g_j(m_j) \quad (5)$$

where $g_j()$ represents some utility function assigned to flow j. We assume that $g_j()$ is non-decreasing and concave. A linear function may be a reasonable candidate. Alternatively, we might choose a form of $g_j()$ that accounts for sampling errors, in which case $g_j()$ will also be a strictly non-decreasing concave function [17, 11]. Also, we can express $u_j(\{m_{ij}\})$ in either of two ways:

$$u_j = g_j(1 - \prod_i (1 - m_{ij})) \quad (6)$$
$$u_j = g_j(\sum_i m_{ij}) \quad (7)$$

Equation (6) models a monitor that independently samples packets. Equation (7) models a system where a monitor can mark sampled packets and sample only unmarked packets.

3 A set of passive monitoring problems without sampling

In this section, we introduce several monitoring problems without sampling; i.e., under the assumption that each monitor collects *all* packets of monitored flows, i.e., $m_{ij} = 1$ or 0 for all i, j, and that there are no routing changes. We will investigate these monitoring problems when routing changes are due to a single link failure in Section 6. We consider the following three problems — the Budget constrained maximum coverage problem (BCMCP), the Maximum deployment cost problem (MDCP), and the Minimum deployment and operating cost problem (MDOCP). We show that all of these problems are NP-hard and that there exist approximate algorithms that give results close to optimum. Table 2 summarizes the set of passive monitoring problems described in this section.

3.1 Budget Constrained Maximum coverage problem without sampling (BCMCP)

The objective of BCMCP is to maximize the monitoring reward subject to a constraint on the total deployment cost. In our initial formulation, we ignore the operating cost. Once a monitor is deployed at a link, all flows carried by the link are fully monitored. We maximize monitoring coverage through an appropriate assignment of monitors to links.

Let B denote the budget constraint, and variable x_j indicate whether flow j is monitored, where $x_j = 1$ means that flow j is monitored and $x_j = 0$ means that flow j is not monitored [2]. The BCMCP can be formulated as an integer linear program:

$$\begin{aligned}
\text{Maximize } & \sum_{j\in D} v_j(x_j) \\
\text{subject to } \quad & x_j \le \sum_{i: j\in S_i} y_i, \quad j \in D \\
& \sum_{i\in L} f_i y_i \le B \\
& y_i \in \{0,1\}, \quad i \in L \\
& x_j \in \{0,1\}, \quad j \in D
\end{aligned}$$

This problem can be shown to be NP-hard by a straightforward reduction from the budgeted maximum coverage problem (MCP). The budgeted maximum coverage problem is defined as follows. We define a collection of sets $Z = \{Z_1, Z_2, \ldots, Z_m\}$ with associated costs $\{c_i\}_{i=1}^m$ over a domain of elements $X = \{x_1, x_2, \ldots, x_n\}$ with associated weights $\{w_j\}_{j=1}^n$. The objective is to determine a collection of sets $Z' \subseteq Z$, such that the total weight of elements covered by Z' is maximized, while the total cost of elements in Z' is less than a given budget K [19]. Here we only provide a sketch of the reduction. Readers are referred to the appendix for a complete proof of the reduction. Z maps to the set of links; the associ-

[2] Here, we use x_j instead of m_{ij} to emphasize the fact that it only takes either 0 or 1. Also, the index $i \in L$ is dropped because the location does not affect the objective function to be maximized.

Table 2
A set of passive flow monitoring problems without sampling

Problem	Placement Cost	Operating Cost	Budget limit	Optimization Goal	Reducible Problem
BCMCP	variable	N/C	yes	max reward of flows	budgeted MCP
MDCP	variable	N/C	no	min placement cost	weighted minimum set cover
MDOCP	variable	variable	no	min placement+operating cost	uncapacitated FLP

ated cost c_i is mapped to the deployment cost f_i; and Z_i maps to the set of flows, S_i, carried by link i. Elements of X, x_j, map to each flow j in a given network. Weight w_j of each element maps to the utility $v_j(x_j)$ of each flow j and the budget K is mapped to the budget constraint for the total deployment cost, B.

We propose a $(1 - 1/e)$-approximation algorithm for BCMCP, which is adapted from the greedy heuristic for the budgeted MCP proposed by S. Khuller et al. [19]. For a special case of BCMCP, where each link has unit deployment cost, the simple greedy heuristic picks, at each step, a link i, with an associated S_i, that maximizes the utility value of the newly covered flows. For the general case in which each link is associated with a different deployment cost, the proposed approximation algorithm computes two candidate solution sets and outputs the one with higher utility value as a final solution. The first candidate solution is the highest utility solution among all subsets of S of cardinality less than k that have cost at most B. Note that k is a tunable parameter that is chosen by a user of the approximation algorithm. The second candidate solution is obtained using simple greedy heuristics with different seed elements. More specifically, for each subset $G \subseteq S$ of cardinality k that has cost at most B, G is included as a very first set of elements to a candidate solution. After that, the candidate solution is augmented with other subsets in S, which greedily maximize the utility value of the newly covered flows $j \in D$. Among candidate solutions, the one with highest utility value is returned as a final greedy solution. The formal description of the approximation algorithm for the general case is presented in Figure 1, where $w(G)$ and $c(G)$ denote the utility value of the flows covered by any set in G and the deployment cost of monitors of G. In addition, we use w_i to denote the sum of the utilities of the flows covered by set S_i, but not covered by any set in G. For $3 \leq k \leq |L|$, the approximation algorithm in Figure 1 is proved to achieve an approximation factor of $(1 - 1/e)$ for BCMCP [19]. The time complexity of the approximation algorithm is $O(|S|^{k+2}|D|)$, where $3 \leq k \leq |L|$.

Approximation algorithm (B,k)

$H_1 \leftarrow$ argmax $\{w(G)$ **, such that** $G \subseteq S$, $|G| \leq k$, **and** $c(G) \leq B\}$
$H_2 \leftarrow 0$
For all $G \subseteq S$, **such that** $|G| = k$ **and** $c(G) \leq B$ **do**
    $U \leftarrow S \backslash G$
    Repeat
        Select link i ($S_i \in U$) **that maximizes** w_i/c_i
        If $c(G) + c_i \leq B$ **then**
            $G \leftarrow G \cup S_i$
        $U \leftarrow U \backslash S_i$
    Until $U = 0$
    If $w(G) > w(H_2)$ **then** $H_2 \leftarrow G$
If $w(H_1) > w(H_2)$, **output** H_1, **otherwise**, **output** H_2

Fig. 1. $(1-1/e)$-approximation algorithm for the general case of BCMCP

3.2 *Minimum deployment cost problem without sampling (MDCP)*

The dual of BCMCP is the minimum deployment cost problem, whose objective is to minimize the placement cost of monitors subject to a minimum monitoring reward. Again, operating costs are not considered, and once a monitor is deployed at a link, all flows carried by the link are captured by the monitor. For example, if the utility of each flow is a constant, MDCP simply minimizes the deployment cost while the number of flows being monitored is equal to or greater than the given monitoring reward, K. If the utility of each flow equals the flows rate, MDCP is defined to minimize the deployment cost subject to the constraint that the sum of rates of monitored flows be equal to or higher than some given rate.

The integer linear program formulation of MDCP is as follows. We want to find an assignment $\{y_i\}$, that:

$$\text{Minimize} \quad \sum_{i \in L} f_i y_i,$$

subject to

$$x_j \leq \sum_{i: j \in S_i} y_i, \quad j \in D$$

$$\sum_{j \in D} v_j(x_j) \geq K$$

$$y_i \in \{0,1\}, \quad i \in L$$

$$x_j \in \{0,1\}, \quad j \in D$$

Since MDCP is the dual problem of BCMCP, it is also NP-hard. We can get an $(\log(|D|)+1)$-approximation solution by adapting the approximation algorithm for the partial K-set cover problem proposed in [25]. We present the approximation algorithm for MDCP in Figure 2, where f_j denotes the monitor deployment cost at link j. In the first step of the algorithm, we normalize the utility values of monitored flows in MDCP because the algorithm requires that all utility values be equal. Without loss of generality, we assume that the utility values of flows take integer values. For each $e_j \in S_i$, if $v_j(x_j) = M$ then replace e_j in S_i with M flows $e_{j1}, e_{j2}, \ldots$, and e_{jM} where $v_{jm}(x_{jm}) = 1$, $m = 1, \ldots, M$. For example, take $S_i = \{e_1, e_2, e_3\}$, where $v_1(x_1) = 3, v_2(x_2) = 2$, and $v_3(x_3) = 1$. Then, replace the original elements $\{e_1, e_2, e_3\}$ with new elements with unit utility values, and define the new S_i as $S_i = \{e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{31}\}$. The key idea of the rest of the algorithm is to choose a subset in each step, taking into account both the number of uncovered flows in D and also the number of uncovered flows in each link i in order to monitor at least K flows. The time complexity of the approximation algorithm is $O(|S|^2|D|)$.

Approximation algorithm (K)

(0) **Replace** each flow $j \in S_i$, where $v_j(x_j) = m$, with new flows $j_1, j_2, \ldots, j_m$, with unit utility values.

(1) **Set** J as the collection of the indices of all S_i

(2) **Set** $J^* = 0$ and $D = \cup_i S_i$

(3) **Set** $r = K - |\cup_{j \in J^*} S_j|$, i.e., r is the number of flows of U yet to be covered in order to obtain a K-cover.

(4) **If** $r \leq 0$, **then** STOP **and** output J^*.

(5) **Find** $i \in J \backslash J^*$ that minimizes the quotient $\frac{f_j}{min(r, |S_j|)}$, for $j \in J \backslash J^*$ and $S_j \neq 0$. In case of tie, take the smallest such i.

(6) **Add** i to J^*. For each $j \in J \backslash J^*$, **set** $S_j = S_j \backslash S_i$. **Set** $D = D \backslash S_i$. **Return** to Step (2).

Fig. 2. $(1 + \log|D|)$-approximation algorithm for the general case of MDCP with integer utility values

3.3 *Minimum deployment and operating cost problem without sampling (MDOCP)*

The objective of MDOCP is to minimize the sum of deployment and operating costs. We assume that the operating cost of a deployed monitoring station is determined by two factors: the average rate of the flow being monitored and the speed of the link where the monitor is deployed. Unlike BCMCP and MDCP, each monitor is allowed to *selectively* monitor a subset of all flows, rather than monitoring all flows. However, the monitoring of a flow cannot be split between different monitoring locations. If a flow is monitored, all packets in that flow are sampled. In this problem setting, we minimize the total deployment and operating cost.

This problem can be formulated as the following integer program. We want to find an assignment to the variables y_i and x_{ij}, such that the following objective function is minimized:

$$\text{Minimize} \quad \sum_{i \in L} f_i y_i + \sum_{i \in L} y_i c_i \sum_{j \in D} \rho_j x_{ij}$$

subject to

$$x_{ij} \leq y_i, \quad i \in L, j \in S_i$$

$$\sum_{i \in L} x_{ij} = 1, \quad j \in D$$

$$x_{ij} \in \{0,1\}, \quad i \in L, j \in S_i$$

$$y_i \in \{0,1\}, \quad i \in L$$

It can be shown that this problem is NP-hard by directly mapping this problem to the well-known uncapacitated facility location problem (FLP). The uncapacitated FLP is defined as follows. We are given a set of locations $N = \{1, \ldots, n\}$ with the inter-location distances denoted as c_{ij}, i, j=$1, \ldots, n$. A facility can be placed at potential facility locations, $F \subseteq N$; building a facility at location $i \in F$ has an associated non-negative cost f_i. There also exist a set of demand points that must be assigned to an open facility, denoted as $D \subseteq N$; for each demand point $j \in D$, we have a positive integral demand d_j that must be shipped to its assigned location. The cost of assigning location i to an open facility at j is c_{ij}

per unit of demand shipped. These costs are assumed to be non-negative, symmetric, and satisfy the triangle inequality; that is, $c_{ij} = c_{ji}$ for all $i, j \in N$, and $c_{ij} + c_{jk} \geq c_{ik}$ for all $i, j, k \in N$. The objective is to determine the set of locations to open facilities and an assignment of demand to the opened facilities, in order to minimize the total cost, i.e., the sum of facility opening cost and the total shipping cost [24, 9]. D. Shmoys et al. proposed a polynomial-time approximation algorithm that finds a solution within a factor of $(1 + 2/e)$ of the optimal, where $1 + 2/e \approx 1.736$. The approximation solution is obtained by rounding an optimal fractional solution to a linear programming relaxation [9].

We can map a MDOCP problem to an uncapacitated facility location problem in the following way. Let's suppose that we have a set of flows M and a set of links L in the network. Let $N = M \bigcup$ L, where N is a set of locations in FLP. Since monitoring stations can be deployed only on links, $F = M$ and $D = L$, where F and D are subsets of locations in FLP. Although the original FLP problem definition requires symmetry and the triangle inequality properties, these are not of concern to us because F and D are disjoint in our special case. The deployment cost of a facility, f_i, is defined as the deployment cost of monitor at link i; the demand d_j is defined as the average rate (i.e., monitoring demand) of flow j. The distance c_{ij} is defined as follows. If flow j traverses link i, then c_{ij} is defined as the unit operating cost of monitor at link i, c_i; otherwise, $c_{ij} = \infty$.

4 A passive monitoring problem with sampling

In this section, we define a monitoring problem under the assumption that a monitor can selectively sample packets in a flow, static routing (i.e., no changes in routing), and allowing the rate for each flow at each monitoring station to be independently adjustable. We will investigate monitoring problems in the presence of routing changes caused by single link failures in Section 6. We introduce the Budget Constrained Maximum Coverage Problem with sampling (BCMCP-S) with the objective of maximizing the monitoring reward given budget constraints. This is the BCMCP problem considered in Section 3, with sampling now considered.

The objective of BCMCP-S is to maximize the total utility of fractional flows being sampled without violating constraints on the monitors' deployment and operating costs. Once a monitor is deployed on a link, a subset of the flows carried by the link can be sampled by the monitor, with the sampling rates of the flows at each monitor chosen independently. Our goal is to maximize the sum of the utilities of the monitored fractional flow, by selecting the number of monitors, their locations, and their sampling rates.

A mixed-integer non-linear program (MINLP) formulation of BCMCP-S is presented below, where S_i represents the set of flows carried by link i, B_1 represents the budget for deployment cost, and B_2 indicates the budget for operating cost. We now want to find an assignment to the variables y_i and m_{ij}, to:

$$\begin{aligned}
\text{Maximize} \quad & \sum_{j \in D} u_j(\{m_{ij}\}) \\
\text{subject to} \quad & \\
& \sum_{i \in L} f_i y_i \leq B_1 \\
& \sum_{i \in L} y_i c_i \sum_{j \in D} \rho_j m_{ij} \leq B_2 \\
& m_{ij} \leq y_i, i \in L, j \in S_i \\
& m_{ij} = 0, \ \ i \in L, j \notin S_i \\
& y_i \in \{0, 1\}, i \in L \\
& m_{ij} \in [0, 1], i \in L, j \in S_i
\end{aligned}$$

In general, it is difficult to directly obtain an optimal solution, because the y_i's are integer variables and $u_j()$ may be nonlinear [1]. For example, if we assume that monitors sample flows independently of each other, the objective function can be represented as $\sum_{j \in D}(g_j(1 - \prod_{i \in L}(1 - m_{ij})))$. To obtain an optimal solution, we can apply the branch-and-bound algorithm combined with gradient projection method to avoid a combinatorial explosion [1]. In practice, a more efficient method is still needed to compute a solution since the branch-and-bound method cannot guarantee a computation time that is polynomial in the number of monitors. Here, we present an approximate solution. We solve BCMCP-S approximately via a two-stage algorithm consisting of a greedy algorithm for integer variables followed by a gradient projection method for the non-integer variables. We present the approximation algorithm in Figure 3. First, we apply a greedy algorithm to obtain an assignment of the y_i variables. The greedy algorithm is similar to the algorithm for BCMCP. Once values are assigned to y_i's by the greedy algorithm, we solve the reduced problem using a gradient projection method, since all of constraints are linear. Since the constraints are linear and the objective function is also assumed to be concave, the iterative solution from the gradient projection method converges to an optimal solution of the reduced problem [3]. However, this solution is an approximate solution of the original problem.

Approximation algorithm (B_1,B_2, k)
    /* Stage 1 */
(1) $H_1 \leftarrow$ argmax $\{w(G)$ **, such that** $G \subseteq S$, $|G| \leq k$,
    and $c(G) \leq B_1\}$
(2) $H_2 \leftarrow 0$
(3) **For all** $G \subseteq S$, **such that** $|G| = k$ **and** $c(G) \leq B_1$ **do**
(4) $U \leftarrow S \backslash G$
(5)   **Repeat**
(6)   **Select** link i ($S_i \in U$) **that maximizes**
        $\sum_{j \in S_i} (u_j(\{x_{ij}\})|_{x_{ij}=1}/\rho_j)/c_i$
(7)   **If** $c(G) + c_i \leq B_1$ **then**
(8) $G \leftarrow G \cup S_i$
(9) $U \leftarrow U \backslash S_i$
(10)   **Until** $U = 0$
(11)   **If** $w(G) > w(H_2)$ **then** $H_2 \leftarrow G$
(12) **If** $w(H_1) > w(H_2)$, **then** $H \leftarrow H_1$,
    otherwise, $H \leftarrow H_2$
(13) **For all** S_i, **if** $S_i \in H$, **then** $y_i \leftarrow 1$,
    otherwise, $y_i \leftarrow 0$
    /* Stage 2 */
(14) **Run** gradient projection method for the reduced
    problem with y_i's and B_2, and output the result

Fig. 3. Two-stage approximation algorithm for BCMCP-S

[3] By changing the sign of the concave objective function in the maximization problem, the problem becomes a minimization problem. Let us denote the new convex objective function with changed sign as f. If f is a convex function, then a local minimum of f over X is a global minimum. If in addition f is strictly convex over X, then there exists at most one global minimum of f over X [5].

5 Evaluation of greedy heuristics, coverage, and marginal gain with additional monitoring points

In this section, we evaluate the effectiveness of our approximation algorithms for BCMCP and BCMCP-S with respect to the optimal solutions. We also investigate the problem of determining the number of monitors needed to achieve a specific level of monitoring reward. Since BCMCP is the dual of MDCP, it suffices to obtain approximation solutions for BCMCP to determine the number of monitors. We first describe the specific parameter settings for the two problems and then show the results of our evaluation.

5.1 Simulation parameter settings

5.1.1 Network topology, traffic matrix, and routing settings

We use both synthetic network topologies and an ISP topology for our study. More specifically, our synthetic topologies, consist of random topologies and Transit-Stub topologies generated by GT-ITM [14]. We use the PoP-level topology of Cable&Wireless, as inferred by the Rocketfuel project [26, 21]. Unfortunately, neither the GT-ITM nor the Rocketfuel dataset provide traffic demand matrices for each topology generated or inferred. Therefore, in order to generate a traffic matrix, we use the technique proposed in [12].

In [12], a synthetic topology is produced using GT-ITM. The original topology model of GT-ITM places nodes in a unit square, thus generating a distance $\delta(x, y)$ between each pair of nodes. These distances result in a random distribution of 2-level graphs with local access arcs and long distance arcs. The topology model does not include a model for the demands; they are modeled as follows. For each node x, two random numbers are chosen: O_x, $D_x \in [0, 1]$. Further, for each pair of nodes (x, y), a random number $C_{(x,y)} \in [0, 1]$ is chosen. The demand between x and y is then $\alpha O_x D_y C_{(x,y)} e^{-\delta(x,y)/2\Delta}$, where the Euclidean distance between x and y is $\delta(x, y)$, Δ is the largest Euclidean distance between any pair of nodes, and α is a parameter that scales the demand [12].

We take a flow to be an ingress-router-to-egress-router flow, unless stated otherwise. However, other flow definitions result in a similar problem formulation. We assume that the path traversed by each flow is determined by a shortest path routing algorithm.

5.1.2 *Utility functions and cost assignment*

Table 3 lists the utility function $v_j()$ or $u_j()$ for each flow j, the deployment cost f_i for each monitor at link i, and the values for all parameters. In Table 3, the traffic demand of flow j, $\rho_j x_j$, is taken as the utility $v_j(x_j)$ for BCMCP. We take the fraction of monitored packets in flow j, m_j, as the utility $u_j(\{m_{ij}\})$, which we will refer to as a linear utility, for BCMCP-S. Specifically, we assume that each monitor samples packets independently in BCMCP-S, such that $m_j = (1 - \prod_i (1 - m_{ij}))$. We take the deployment costs, f_j, to be one, $f_j = 1$.

5.2 *Simulation results of BCMCP*

Figure 4 plots the fraction of monitored packets as a function of the number of monitors, using the greedy BCMCP heuristic for the random network of 10 routers listed in Table 4. Although BCMCP is NP-hard, we can compute optimal solutions for small problem sizes and compare them to solutions produced by our heuristics. The random selection solution shown in Figure 4 represents the fraction of monitored packets when locations chosen randomly with equal probability. For the random selection algorithm, we plot the average over the independent runs. Figure 4 shows that the random selection method performs poorly compared to our greedy BCMCP heuristic. It also shows the effectiveness of the greedy solution: the difference between the optimal and heuristic solutions are quite small. We also observe that by deploying monitors at 29% of the possible locations (in this case, 4 monitors), we can monitor more than 90% of the network's packets. We also observe that the marginal increase in the fraction of monitored packets decreases as additional monitors are added.

Figure 5 shows similar results for the case of a transit-stub topology of 27 nodes listed in Table 4. Because of the combinatorial explosion in the number of links, we could only compute optimal solutions for up to 5 monitors for this problem. However, we also calculated linear program (LP) solutions. When we relax the integer constraints, the BCMCP becomes a linear program (LP), which yields an upper bound. We term the LP solution as "LP bound". If our approximation solution happens to be very close to LP bound, this indicates that it is very close to optimal. Indeed, our approximate solution is very close to optimal, as illustrated in Figure 5. Interestingly, the greedy solution for three monitors results in approximately 90% of the packets being monitored, and produces the same result as the optimal solution. We again observe decreasing marginal returns as the number of monitors increases. In Section 3, we have shown that $(1 - 1/e)$ is the theoretical bound of the approximation ratio of the greedy solution to the optimal solution in BCMCP. However, in Figure 4 and 5, we observe that the greedy solution is actually much closer to optimal. We conjecture that this is due to

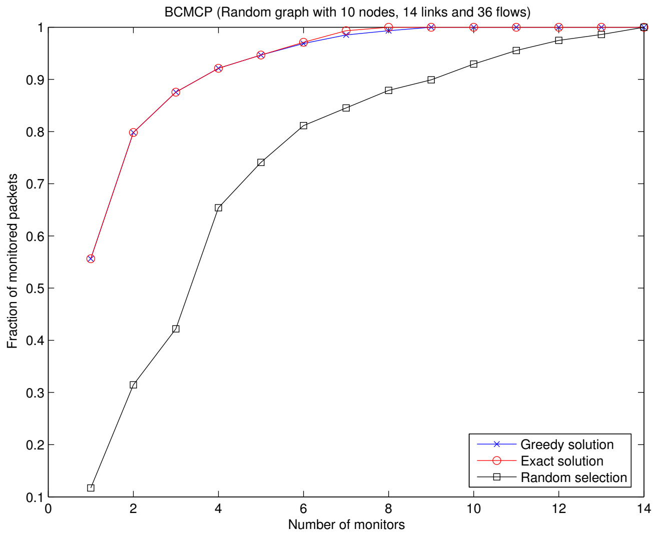

Fig. 4. BCMCP in a 10 node, 14 link, and 36 flow GT-ITM Random topology

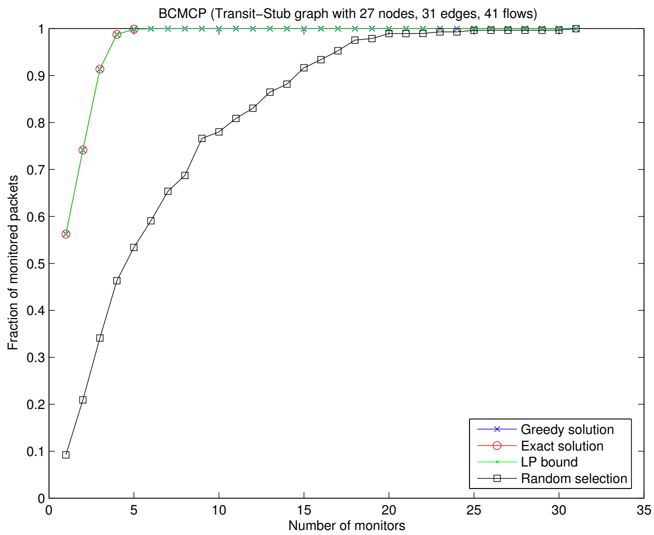

Fig. 5. BCMCP in a 27 node, 31 link, and 41 flow GT-ITM Transit-Stub topology

Table 3
Parameter settings

Parameter	value
f_i	1
ρ_j	traffic demand of flow j
c_i	1
m_j	$(1 - \prod_i (1 - m_{ij}))$
$v_j(x_j)$	$\rho_j * x_j$
$u_j(\{m_{ij}\})$	m_j

Table 4
GT-ITM topologies and the total number of flows

Generation method	Num of nodes	Num of links	Num of flows	Additional information
Random	10	14	36	
Transit-stub	27	31	41	3 trans-node; 2 stubs/trans-node; 4 nodes/stub
Transit-stub	100	187	8885	4 trans-nodes; 3 stubs/trans-node; 8 nodes/stub

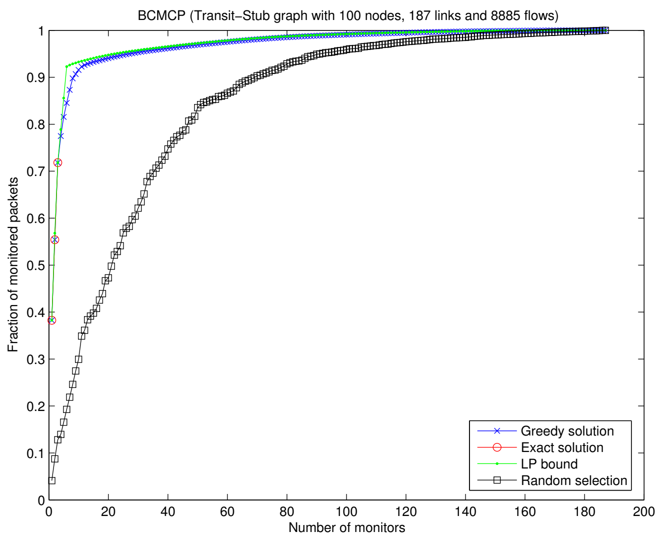

Fig. 6. BCMCP in a 100 node, 187 link, and 8885 flow GT-ITM Transit-Stub topology

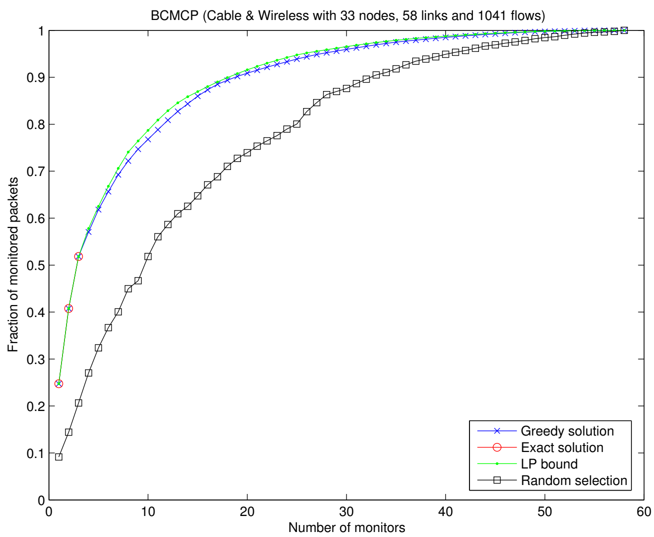

Fig. 7. BCMCP in Cable&Wireless PoP-level topology from Rocketfuel dataset

various factors such as shortest-path routing, the hierarchical structure of the topologies, and all-pair traffic demands among nodes. Another interesting observation is that in Figure 5, a smaller fraction of monitoring locations is needed than in Figure 4 to achieve similar monitoring rewards. For example, to monitor 90% of the packets, only 10% of the locations require monitoring in the case of the random stub network, while, about 29% of the locations require monitoring in the case of the transit-stub network. We conjecture that this is because the transit-stub model produces transit links and that these links are traversed by most of the flows.

Figure 6 shows the results for a larger transit-stub topology of 100 nodes. Again we could not compute exact solutions for all possible number of monitors. However, we can again compare the greedy solution with the LP bound, which indicates that the greedy and optimal solutions are indeed very close. Figure 6 also exhibits a decreasing monitoring reward gain. In addition, a smaller fraction of monitoring locations is required in the 100 node network of Figure 6 than by the small 27 node network to achieve a same level of monitoring reward (up to 90% of monitored packets), suggesting that a larger topology tends to require a smaller fraction of monitoring locations for the same monitoring reward.

In Figure 7, we evaluate our BCMCP heuristic us-

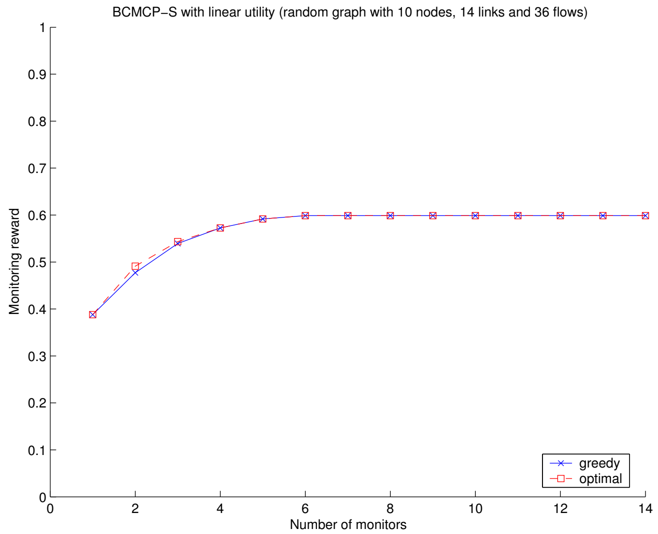

Fig. 8. BCMCP-S in a 10 node, 14 link, and 36 flow GT-ITM Random topology

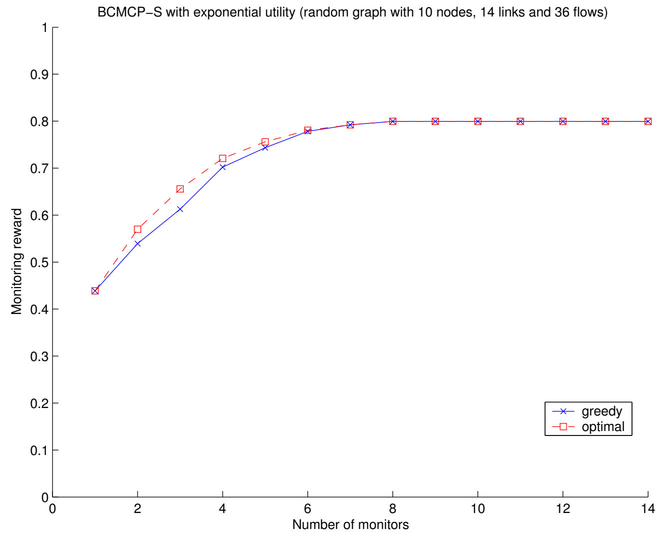

Fig. 9. BCMCP-S in a 10 node, 14 link, and 36 flow GT-ITM Random topology

ing the PoP-level topology [4] of Cable & Wireless as inferred by the Rocketfuel project [26, 21]. Since the topology data from the Rocketfuel project does not contain traffic matrices, we again generate the traffic demand matrix according to the model in [12]; a shortest path routing is again used to find the routing path for each flow, assuming that a single path is always used. In Figure 7, we again observe similar trends as in Figures 4, 5, and 6 such as a diminishing reward gain as the number of monitors increases. Furthermore, the comparison of the greedy solution to the LP bound again indicates the high quality of the greedy solution. On the other hand, Figure 7 shows the poor quality of a random allocation.

5.3 Simulation results for BCMCP-S

Figure 8 shows the heuristic and optimal solutions to BCMCP-S for the 10-node random graph listed in Table 4. In addition to the parameter settings in Table 3, we select a budget, B_2, for the total operating cost such that it is impossible to sample 100% of packets of all flows in the 10 node network of Figure 8 [5]. We generate optimal solutions in a brute-force way by selecting every set of monitors that satisfies the budget constraint, B_1, and running the gradient projection method for each selected set of monitors. The graph shows that the greedy solution is quite close optimal. In Figure 9, we use the same random topology as in Figure 8, but replace the linear utility with an exponential utility function, $u_j(\{m_{ij}\}) = (1 - \exp(-5.0 * m_j))$, where $m_j = (1 - \prod_i (1 - m_{ij}))$. We again observe that the greedy solution is close to optimal. In addition, both Figure 8 and 9 show again that the marginal increase in monitoring reward decreases as additional monitors are added. Although we omit the exact number of sampled flows and sampling rates here, we observe that in the case of an exponential function, more flows are monitored (although lower sampling rates are obtained on average) than in the former case that uses the linear function.

6 Discussion and extensions

In this section, we extend the monitoring problems proposed in Section 3 and 4, to allow for the random arrival and departure of the flows and to account for routing changes caused by single link failure. We find that the resulting problems are essentially the same as the original problems in terms of complexity and approximation algorithms.

In the previous sections, all monitoring problems were formulated under the assumption that the number of flows between each origin and destination (OD)

[4] We treat each POP as a single router in this POP-level topology for our evaluation.

[5] Specifically, when 100% of packets of all flows in the 10 node network is sampled, the minimum total operating cost is 2492. Here, we select 500 (out of 2492) as a budget constraint, B_2, for the total operating cost.

pair is long-lived. We now consider the dynamic arrival/departure of flows between an OD pair in a network. In order to consider a varying number of flows in the monitoring problems, we can simply take each OD pair as a basic monitoring unit instead of a flow. Accordingly, we take the average aggregate rate of flows between an OD pair as the average OD pair rate. Although we replace flows and flow rate with OD pairs and OD pair rate respectively in those monitoring problems, the complexity and approximation algorithm for the new problems do not change.

We next consider the traffic monitoring problems in the face of single link failure and subsequent routing changes caused by the link failure. We assume that the sampling rate of a flow can be changed when a route changes, but that the locations of the monitors remain fixed in the network. In order to describe the resulting monitoring problems, we extend the set of parameters defined in Section 2. We summarize the additional parameters in Table 5. We add the superscript (k) to a quantity to denote that quantity in the presence of the failure of link $k \in \{1, \ldots, |L|\}$. When a link fails in the network, the failure is usually detected by all routers and their routing tables are updated to set up new routing paths [6]. Clearly, each flow may take up to additional $|L|$ different routing paths determined by the value of k. $S_i^{(k)}$ denotes the set of flows traversing link i in the failure of link k. Also, the probability that link k is in failure is denoted by the parameter $p^{(k)}$ [7].

- **BCMCP-LF (Budget Constrained Maximum Coverage Problem considering single Link Failure):**
  The integer linear program formulation of BCMCP-LF is as follows. We want to find an assignment for the variables y_i, in order to:

$$
\begin{array}{lll}
\text{Maximize} & \multicolumn{2}{l}{\sum_{j \in D} [\sum_{k=1}^{|L|} p^{(k)} v_j(x_j^{(k)})] + \sum_{j \in D} (\prod_{k=1}^{|L|} (1 - p^{(k)})) v_j(x_j)} \\
\text{subject to} & x_j^{(k)} & \leq \sum_{i: j \in S_i^{(k)}} y_i, \;\; j \in D \\
 & x_j & \leq \sum_{i: j \in S_i} y_i, \;\; j \in D \\
 & \sum_{i \in L} f_i y_i & \leq B \\
 & y_i & \in \{0,1\}, \;\; i \in L \\
 & x_j^{(k)} & \in \{0,1\}, \;\; j \in D \\
 & x_j & \in \{0,1\}, \;\; j \in D \\
 & k & \in \{1, \ldots, |L|\}
\end{array}
$$

It can be shown that the a BCMCP-LF can be transformed to a BCMCP problem in linear time. Here we sketch the transformation. Given a BCMCP-LF problem, we replicate each flow j, $|L|$ times such that we have $(|L|)$ new flows, $j^{(1)}, \ldots, j^{(|L|)}$, which may take different routing paths. We expand the S_i's for each i where $i \in L$, such that $j^{(k)} \in S_i, j \in L, k = \{1, \ldots, |L|\}$. For example, $j^{(k)} \in S_i$ indicates that original flow j traverses link i in the case of link k failure. Clearly, by treating each replicated flow created from original flow j independently from each other and by assigning $p^{(k)} v_j(x_j^{(k)})$ to the utility value of flow $j^{(k)}$, we get a BCMCP problem. Therefore, we can simply apply our algorithm for BCMCP to BCMCP-LF.

- **MDCP-LF (Minimum Deployment Cost Problem considering single Link Failure):**
  The integer linear program formulation of MDCP-LF considering single link failures is as follows. We want to find an assignment for the variables y_i, in order to:

$$
\begin{array}{lll}
\text{Minimize} & \sum_{i \in L} f_i y_i, & \\
\text{subject to} & & \\
 & x_j^{(k)} & \leq \sum_{i: j \in S_i^{(k)}} y_i, \;\; j \in D \\
 & x_j & \leq \sum_{i: j \in S_i} y_i, \;\; j \in D \\
 & \multicolumn{2}{l}{\sum_{j \in D} [\sum_{k=1}^{|L|} (p^{(k)} v_j(x_j^{(k)})) +}
\end{array}
$$

[6] Here, we assume that at least an alternative routing path exists for any single link failure, since ISPs usually design their networks such that they are not partitioned even in the event of large-scale failure involving multiple links and/or nodes [16].

[7] We do not present MDOCP-LF problem here. We believe this problem is rather unrealistic because we might end up with deploying monitors at most of the links to cover all flows regardless of routing changes.

Table 5
Additional parameters in the extended model considering single link failures.

Parameter	Definition
$x_j^{(k)}$	$\{0,1\}$ indicator if flow j is being monitored, $j \in D$, in case of link k failure
$m_{ij}^{(k)}$	Fraction of flow j sampled by a monitor at link i, $i \in L$, $j \in D$, in case of link k failure
$m_j^{(k)}$	Fraction of flow j sampled by all monitors in case of link k failure
$u_j(\{m_{ij}^{(k)}\})$	Nondecreasing concave function of the fraction of flow j sampled by all monitors in case of link k failure
$v_j(x_j^{(k)})$	Nondecreasing concave function of whether flow j is fully captured by any monitor in case of link k failure
$p^{(k)}$	Probability that link k is in failure

$$\prod_{k=1}^{|L|}(1-p^{(k)})v_j(x_j)] \quad \geq K$$

$$y_i \quad \in \{0,1\}, \quad i \in L$$

$$x_j^{(k)} \quad \in \{0,1\}, \quad j \in D$$

$$x_j \quad \in \{0,1\}, \quad j \in D$$

$$k \quad \in \{1,\ldots,|L|\}$$

Similar to BCMCP-LF, we first transform a given MDCP-LF problem to a MDCP problem and then apply the approximation algorithm for MDCP to the transformed problem.

- **BCMCP-S-LF (Budget Constraint Maximum Coverage Problem considering single link failure with sampling):** The mixed-integer non-linear program formulation of extended BCMCP-S-LF considering single link failure is as follows. We want to find an assignment to the variables y_i and $m_{ij}^{(k)}$, in order to:

$$\text{Maximize} \quad \sum_{j\in D}[\sum_{k=1}^{|L|} p^{(k)}u_j(\{m_{ij}^{(k)}\})] + \sum_{j\in D}(\prod_{k=1}^{|L|}(1-p^{(k)}))u_j(\{m_{ij}\})$$

subject to

$$\sum_{i\in L} f_i y_i \quad \leq B_1$$

$$\sum_{i\in L} y_i c_i \sum_{j\in D} \rho_j [\sum_{k=1}^{|L|} p^{(k)} m_{ij}^{(k)} + \prod_{k=1}^{|L|}(1-p^{(k)})m_{ij}] \quad \leq B_2$$

$$m_{ij}^{(k)} \quad \leq y_i, i \in L, j \in S_i^{(k)}$$

$$m_{ij}^{(k)} \quad = 0, \quad i \in L, j \notin S_i^{(k)}, j \in D$$

$$m_{ij} \quad \leq y_i, i \in L, j \in S_i$$

$$m_{ij} \quad = 0, \quad i \in L, j \notin S_i, j \in D$$

$$y_i \quad \in \{0,1\}, i \in L$$

$$m_{ij}^{(k)} \quad \in [0,1], i \in L, j \in S_i^{(k)}$$

$$m_{ij} \quad \in [0,1], i \in L, j \in S_i$$

$$k \quad \in \{1,\ldots,|L|\}$$

It can be shown that the algorithm for BCMCP-S-LF is same as the algorithm for BCMCP-S. The key idea is also to transform a BCMCP-S-LF problem to a BCMCP-S problem. Here we sketch the transformation. Given a BCMCP-S problem, we replicate each flow j, by $|L|$ times such that we have $|L|$ new flows, $j^{(1)}, \ldots, j^{(|L|)}$. We expand S_i's for each i where $i \in L$, such that $j^{(k)} \in S_i, j \in L, k = \{1,\ldots,|L|\}$. By treating each replicated flow created from original flow j independently from each other and also by assigning $p^{(k)}u_j(\{m_{ij}^{(k)}\})$ to the utility function of flow $j^{(k)}$ monitored at link i, we get a BCMCP-S problem. Therefore, we can simply apply the algorithm for BCMCP-S to BCMCP-S-LF.

7 Related work

Several recent efforts have addressed the problem of placing *active* monitors and packet filters in networks. In addition, the question of how to sample packets in a single monitor has been addressed by several researchers. The problem formulation presented in this paper is unique when compared to prior work in these areas. We summarize prior work in this section.

7.1 Number and location of tracers in the Internet

Measurement points that send "active probe" messages may be used to obtain topology and performance measurement such as link delay and existence of faults etc. The location of these measurement devices or beacons has been determined by various heuristics in the literature [18, 4, 20, 6]. These efforts are similar to our work in the sense that near optimal locations of measurement devices are obtained using greedy heuristics. Horton et al. [15] determined the minimal number of required beacons in a network and their near-optimal locations. Barford et al. [3] presented empirical observations showing that the marginal utility of adding additional active measurement sites declines rapidly after the second or third site. Both works show that a relatively small number of active measurement points is generally sufficient to obtain an accurate network topology [3, 15]. This conclusion is similar to our observation that only a small number of passive monitors is necessary to achieve high monitoring coverage. Note that these earlier works were concerned with active, rather than passive, monitors. More recently, in parallel with our work, Chua et al. [8] reported their observation that many end-to-end paths share some common links in ISP networks, particularly in Abilene network. Their observation confirms our experimental result with Cable&Wireless PoP-level topology.

7.2 High coverage power with a small number of passive monitoring/filtering locations

K. Park and H. Lee [22] showed that the well-known vertex cover problem can be used to approximate the problem of placing router-based packet filters to prevent distributed DoS Attack. Since the vertex cover problem is known to be NP-hard, they investigated several greedy heuristics. They also argued that the installation of route-based packet filters in the border routers of a small number of ASes is enough to achieve high defense coverage against DDoS attacks, because it is known that AS graphs follow a power-law. Our work is similar to their work in the sense that greedy heuristics are proposed because of the NP hardness of the problems. We also show that a few locations are sufficient to achieve high coverage. However, we deal with the problem of distributed monitoring, a different problem than route-based packet filtering systems. Also, we propose both deployment and operating strategies for distributed monitoring.

7.3 Network tomography

The objective of Minimum cost Multicast Tree Cover Problem (MMTCP) in [2] is: given a set of links whose behavior is of interest, how does one choose a set of minimum cost multicast trees within the network to determine the behavior of the links in question, particularly link loss rate. They introduce a cost function that accounts for a per tree cost and per link costs. The MMTCP is similar to our MDCP because simple greedy heuristics for weighted minimum set-cover problem was proposed for both problems. However, MMTCP is again concerned with active measurements, which is different from a passive measurement problem considered by MDCP. In addition, sampling is not considered in [2].

7.4 Sampling strategies

Sampling methodologies at a single measurement point [7, 17, 10] are related to BCMCP-S. In practice, sampling accuracy may be measured by the magnitude of the variance or the relative size of the confidence interval to the mean value of unbiased estimators. If we model $u_j(\{m_{ij}\})$ as the sampling accuracy of flow j, the objective of BCMCP-S becomes that of maximizing the overall sampling accuracy under a constrained monitoring cost. However, in [7, 17], the minimum number of sampled packets is calculated under the given bounding constraint for the sampling error to infer the total volume of packets with some common attributes. In [10], in order to jointly control the volume of samples $\hat{N}$ and the variance of the estimator $\hat{X}$ without assumptions on the distribution of the sizes x_i, a cost function $C_z(p) = Var\hat{X} + z^2 E\hat{N}$ is introduced and it is shown that a size-dependent sampling which dynamically chooses sampling rate $p_z(x) = min\{1, x/z\}$ [8] minimizes the cost function $C_z(p)$.

8 Conclusion

In this paper, we have presented near-optimal monitor placement and operating strategies in a distributed monitoring system, which operate either

[8] where z is a tunable parameter and x is the size of a given flow being monitored.

in sampling or non-sampling mode. Each deployment strategy determines the maximum number of monitors and their locations under a given budget constraint or determines the minimum deployment cost for a maximum number of monitors. Also, the operating strategy of each monitor determines the flow sampling rate. More specifically, we first introduced novel monitoring cost and reward models for a distributed passive monitoring system, which can accommodate both sampling and non-sampling modes of the monitoring system. Based on these models, we formulated a set of placement and operating problems assuming different constraints for budget, coverage, and routing requirements. We also showed that various placement problems are NP-hard. We proposed approximation algorithms based on greedy heuristics to determine placement locations and used a gradient projection method to get sampling rates.

Secondly, we evaluated the relationship between the number of monitors and the maximum reward of flows using both synthetic and an ISP topologies. Through the experiments, we showed the decreasing gain of monitoring reward whenever additional monitors are added. Also, we showed that only a small fraction of links need be monitored to achieve a high level of monitoring reward.

Finally, we presented the experimental results showing that the proposed approximation algorithms achieve very good solutions. More specifically, according to our experiments with network topologies, our proposed greedy solutions achieved much better approximation ratios than the well-known theoretical bounds for the approximation algorithm for budgeted maximum coverage problem, $O(1 - 1/e)$. We conjecture that the hierarchical structure of topologies, shortest path routing, and all-pair traffic demands among nodes result in a small set of links carrying most of the flows. We conjecture that such links are the early candidate links in our greedy solutions, and are consistent with the links chosen in optimal solutions. However, further study on this issue is needed to validate this insight.

As on-going work, we are evaluating our heuristics on more diverse real ISP topologies such as Sprint and AT&T networks as inferred by Rocketfuel projects. We plan to evaluate the effectiveness of the approximation algorithm for MDOCP. In addition, we are investigating whether a tighter bound of the approximation ratio of the greedy solution can be found for power-law networks.

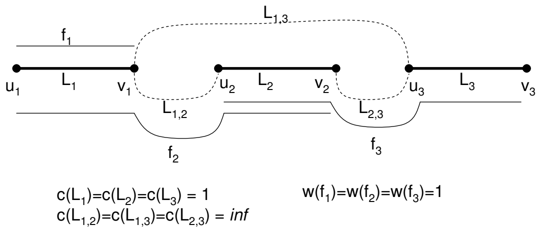

Fig. 10. The graph $G_{R(I)}(V, E)$ and the flow set $F_{R(I)}$ for the given instance of the MCP problem

Appendix

The hardness of the BCMCP problem

In the following, we show that the BCMCP problem is NP-hard by presenting a polynomial reduction from the budgeted maximum coverage problem (MCP) to the BCMCP problem. In an instance of $I(Z, X)$ of the MCP problem, $Z = \{Z_1, Z_2, \ldots, Z_m\}$ is a collection of m sets with associated costs $\{c_i\}_{i=1}^m$ over a domain of n elements $X = \{x_1, x_2, \ldots, x_n\}$ with associated weights $\{w_j\}_{j=1}^n$. The objective is to determine a collection of sets $Z' \subseteq Z$, such that the total weight of elements covered by Z' is maximized, while the total cost of elements in Z' is less than a given budget B [19]. Consider an instance $I(Z, X)$ of the MCP problem. Our reduction $R(I)$ constructs a graph $G_{R(I)}(V, E)$, a set of flows $F_{R(I)}$ on the graph, and a budget constraint for the total deployment cost $K_{R(I)}$. Each edge is associated with a monitoring cost and the graph contains the following nodes and non-auxiliary and auxiliary edges. For each set $Z_i \in Z$, the graph contains two connected nodes u_i and v_i, which form a non-auxiliary edge L_i with monitoring cost c_i. For each pair of v_i and u_j where $i < j$, we connect them with an *auxiliary* edge L_{ij} in the graph. We assign *infinite* monitoring cost to such L_{ij}'s. In addition, for each element $x_j \in X$, we construct a flow f_j which traverses each non-auxiliary edge L_i where $x_j \in Z_i$. The flow f_j also traverses each auxiliary edge L_{i_1,i_2} where $i_1 < i_2$, $x_j \in Z_{i_1}$, and $x_j \in Z_{i_2}$, however, there exists no Z_{i_3} such that $x_j \in Z_{i_3}$ and $i_1 < i_3 < i_2$. Each flow f_j is assigned a weight w_j. Also, the $K_{R(I)}$ is set equal to B. Figure 10 shows an example of the reduction, $G_{R(I)}$ and $F_{R(I)}$ from an instance of MCP problem $I(Z, X)$, where $Z = \{Z_1, Z_2, Z_3\}, X = \{x_1, x_2, x_3\}, Z_1 = \{x_1, x_2\}, Z_2 = \{x_2, x_3\}$, $Z_3 = \{x_3\}$, $w_1 = w_2 = w_3 = 1$ and $c_1 =$

$c_2 = c_3 = 1$.

We claim that there exists a solution of total weight k to the given MCP problem if and only if there exists a solution of total weight k to the BCMCP instance defined by the graph $G_{R(I)}(V, E)$ and the set of flows $F_{R(I)}$. We first show that if there exists a solution to the MCP problem of weight k then there exists a set T of edges that allows us to monitor a set of flows of weight k. Let the solution of the MCP problem $Z' = \{Z_{i_1}, \ldots, Z_{i_l}\}$, such that the total cost of elements in Z' is not larger than a given budget B and the total weight of elements covered by Z' is k. Then, the set T of monitored edges contains $L_{i_1}, \ldots, L_{i_l}$. We show that the set T contains edges of total cost at most B, that cover flows of total weight at least k. Clearly, the total cost of the monitored edges in T is at most B. Each $Z_{i_p} \in Z'$ in the solution of the MCP problem covers all the elements x_j's, where $x_j \in Z_{i_p}$. Because of the 1-to-1 mapping, each L_{i_p} in the BCMCP problem monitors all the flows f_j's traversing the edge L_{i_p}. Since $\sum_{x_j \in \cup_{(Z_{i_p} \in Z')} Z_{i_p}} w(x_j) = k$, the set of edges $\{L_{i_1}, \ldots, L_{i_l}\}$ can monitor flows of total weight k.

Also, we show that if there exists a set of monitored edges of total cost less than B that monitors flows of weight at least k, then there exists a solution for the MCP problem of weight at least k. Let the solution of the BCMCP problem consist of a set of l monitors which monitor flows of weight k. Since the monitoring cost of each auxiliary edge is infinity and all the flows carried by an auxiliary edge $L_{i,j}$ is also carried by both L_i and L_j, the set of l edges must consist of non-auxiliary edges only. More specifically, the solution for the BCMCP problem is composed of $L_{i_1}, \ldots, L_{i_l}$, to monitor flows of weight at least k.

Let us suppose that there is no solution to the MCP problem of weight k. Consequently, there are only two possibilities for $Z' = \{Z_{i_1}, \ldots, Z_{i_l}\}$. Either the weight of the elements in X covered by Z' is less than k or the total cost $\sum_{Z_{i_p} \in Z'} c(Z_{i_p})$ is greater than the constraint B. Again, because of the 1-to-1 mapping, each L_{i_p} in the BCMCP problem monitors all the flow f_j's traversing the edge L_{i_p}. If the total cost is greater than the constraint B in the MCP problem, the total cost of $L_{i_1}, \ldots, L_{i_l}$ should be greater than B in the BCMCP problem. This is a contradiction. If $\sum_{x_j \in \cup_{(Z_{i_l} \in Z')} Z_{i_l}} w(x_j) = k'$ such that $k' < k$, the total weight of flows monitored at edges $L_{i_1}, \ldots, L_{i_l}$ is k'. This is also a contradiction. Therefore, there is a solution to the MCP problem of weight at least k.

Acknowledgment

An earlier version of this paper was also published in the proceedings of IEEE INFOCOM 2005. This research was supported in part by the National Science Foundation under NSF grants EIA-0080119, ANI-0434667, and ANI-0085848. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies. We would like to thank Honggang Zhang, Chun Zhang, and the anonymous reviewers for useful comments and suggestions.

References

[1] C. Adjiman, C. Schweiger, and C. Floudas. *Mixed-integer nonlinear optimization in process synthesis*. Handbook of Combinatorial optimization (D.-Z. Du and P.M Pardalos(Eds.)), 1998.

[2] M. Adler, T. Bu, R. Sitaraman, and D. Towsley. Tree layout for internal network characterizations in multicast networks. In *Proc. International Workshop on Networked Group Communication (NGC)*, 2001.

[3] P. Barford, A. Bestavros, J. Byers, and M. Crovella. On the marginal utility of network topology measurements. In *Proc. ACM SIGCOMM Internet measurement workshop (IMW)*, 2001.

[4] Y. Bejerano and R. Rastogi. Robust monitoring of link delays and faults in IP networks. In *Proc. IEEE INFOCOM*, 2003.

[5] Dimitri P. Bertsekas. *Nonlinear programming*. Athena Scientific, 1999.

[6] Y. Breitbart, C. Chan, and et al. Efficiently monitoring bandwidth and latency in IP networks. In *Proc. IEEE INFOCOM*, 2001.

[7] B. Choi, J. Park, and Z. Zhang. Adaptive random sampling for traffic load measuremnt. In *Proc. ACM SIGMETRICS*, 2002.

[8] D. Chua, E. Kolaczyk, and M. Crovella. Efficient monitoring of end-to-end network properties. In *Proc. IEEE INFOCOM*, 2005.

[9] F. Chudak and D. Shmoys. Improved approximation algorithms for the unpacacitated facility location problem. *ACM SIAM Journal on computing*, (1), 2003.

[10] N. Duffield, C. Lund, and M. Thorup. Learn

more, sample less: control of volume and variance in network measurement. 2002.

[11] N. Duffield, C. Lund, and M. Thorup. Properties and prediction of flow statistics from sampled packet streams. In *Proc. ACM SIGCOMM Internet Measurement Workshop 2002*, 2002.

[12] B. Fortz and M. Thorup. Internet traffic engineering by optimizing OSPF weights. In *Proc. IEEE INFOCOM*, 2000.

[13] C. Fraleigh, C. Diot, B. Lyles, S. Moon, P. Owezarski, K. Papagiannaki, and F. Tobagi. Design and deployment of a passive monitoring infrastructure. In *Proc. of Passive and active measurement workshop*, 2001.

[14] GT-ITM. Georgia tech internetwork topology models. http://www.cs.gatech.edu/projects/gtitm/.

[15] J. Horton and A. Ortiz. On the number of distributed measurement points for network tomography. In *Proc. ACM Internet measurement conference*, 2003.

[16] G. Iannaccone, C. Chuah, S. Bhattacharryya, and C. Diot. Feasibility of IP restoration in a tier-1 backbone. *IEEE Network magazine*, March 2004.

[17] INMON Corp. sFlow accuracy and billing. http://www.inmon.com/pdf/sFlowBilling.pdf, 2001.

[18] S. Jamin, C. Jin, and et al. On the placement of Internet instrumentation. In *Proc. IEEE INFOCOM*, 2000.

[19] S. Khuller, A. Moss, and J. Naor. The budgeted maximum coverage problem. *Information Processing Letters*, 1997.

[20] L. Li, M. Thottan, and et. al. Distributed network monitoring with bounded link utilization in IP networks. In *Proc. IEEE INFOCOM*, 2003.

[21] R. Mahajan, N. Spring, and D. Wetherall. Inferring link weights using end-to-end measurements. In *Proc. ACM Internet Measurement Workshop (IMW)*, 2002.

[22] K. Park and H. Lee. On the effectiveness of route-basd packet filtering for distributed DoS attack prevention in power-law Internets. In *Proc. ACM SIGCOMM*, 2001.

[23] M. Sharma, G. Iannaccone, and S. Bhattacharrya. On the placement of monitoring devices in an IP network. Technical Report RR03-ATL-112424, Sprint ATL, 2004.

[24] D. Shmoys, E. Tardos, and K. Aardal. Approximation algorithms for facility location problems. In *Proc. ACM Symposium on Theory of Computing*, 1997.

[25] P. Slavik. Improved performance of the greedy algorithm for the minimum set cover and minimum partial cover problems. *Information Processing Letters*, 1997.

[26] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP topologies with rocketfuel. In *Proc. ACM SIGCOMM*, 2002.