Locating network monitors: complexity, heuristics, and coverage

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

Department of Computer Science
University of Massachusetts
Amherst, MA 01003
{kwsuh, kurose, towsley}@cs.umass.edu

†The MathWorks, Inc.
Natick, MA 01760
Yang.Guo@mathworks.com

UMass Computer Science Techincal Report 2004-61

July-14-2004

Abstract

There is increasing interest in concurrent passive monitoring of IP flows at multiple locations within an IP network. The common 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 consider minimum cost and maximum coverage problems under various budget constraints. We show that all of the defined problems 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 are often enough to monitor most of the traffic in an entire IP network.

I. 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 [1]. In practice, a monitor is placed inside a router or deployed as a standalone measurement box that taps into a communication link. Once a monitor is placed on a link, it may capture or sample packets carried by the link depending on its specific sampling configuration. In order to observe a large fraction of a network's traffic, we need 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 also factors into its operating

This research was supported in part by the National Science Foundation under NSF grants EIA-0080119, ANI-9973635, 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. The Appendix is included in the submission for the reviewers' perusal and will be excluded from the final version of the paper.

cost. Considering that the per-packet operating cost of each monitor depends mainly on the link speed, we may take advantage of the monitors with lower per-packet operating costs 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 novel monitoring cost and coverage models of both sampling and non-sampling monitor modes for a distributed monitoring system. The monitoring cost is defined as the sum of deployment cost and operating cost. Based on these models, we formulate several variations of the problem of optimizing 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 the greedy solutions provide solutions that are quite close to optimal for a variety of 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 [2–5]. Also, Horton et al. [6] determined the minimal number of required beacons in a network and their optimal locations in order to obtain accurate Internet topology. While these works investigated active monitoring, the IPMon project at Sprint [7] 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. [8] proposed a heuristic for locating passive monitors. However, they do not consider operating costs or 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, the key ideas, and the performance metrics which include deployment cost, operating cost, and monitoring reward. Sections III and IV formulate various budget-constrained and coverage-constrained monitoring problems in both sampling and non-sampling modes of operation along with their complexity analysis. In Section V, we evaluate the proposed greedy heuristics and examine the coverage issue. A summary of related work is presented in Section VI. We conclude with a summary of our results and a discussion of future work.

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, $i \in L$, $j \in D$.
m_{ij}	Fraction of flow j sampled by a monitor at link i
M_j	Fraction of flow j sampled by all monitors
$u_j(M_j)$	Nondecreasing concave function of the fraction of flow j sampled by all monitors

TABLE I
PARAMETERS IN THE MODEL.

II. PROBLEM SETTING

In this section, we consider the distributed monitoring problem, and propose three models — a deployment cost model, an operating cost model, and a monitoring reward model — to represent different aspects of the monitoring problem. We will further investigate the interaction among these three models in Sections III and IV, and show that it is possible to achieve near-optimal monitor placement and operating conditions under various budget constraints and monitoring objectives.

We represent an IP network as an undirected graph, $G(V, E)$, where V and $E \subseteq V$ x 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 network-to-network flow [1], ingress router-to-egress router flow, host-to-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 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 I.

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 *(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 design 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 I). Here, we use D and S_i to denote a set of all flows in the network and a set of all flows carried by link i, respectively. We use y_i to indicate whether a monitor is deployed at link i:

[1] PoP-to-PoP flow, AS-to-AS flow, and Customer-network-to-Customer-network flow may be typical examples of network-to-network flow.

- **Deployment cost model.** The deployment cost model captures the cost associated with deploying a monitoring station. We use f_i to 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:

$$C_D = \sum_{i \in L} f_i y_i \tag{1}$$

- **Operating cost model.** The operating cost model captures the cost associated with the monitor's sampling operation, assuming that a passive monitoring station is able to monitor traffic that traverses link in both directions. We use c_i to denote the unit operating cost of monitor i. This could represent the cost of sampling a single packet at monitor i. The values of c_i at each monitor can differ, e.g., because of link speed. We also use m_{ij} and ρ_j to 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:

$$C_O = \sum_{i \in L} y_i c_i \sum_{j \in D} \rho_j m_{ij} \tag{2}$$

- **Monitoring reward model.** The monitoring reward model captures the benefit of traffic monitoring. Let utility function $u_j(M_j)$ denote the benefit gained by monitoring flow j, where M_j is the fraction of flow j that has been monitored in the network. For example, $u_j(M_j)$ may simply represent the monitored traffic demand of flow j.

$$C_M = \sum_{j \in D} u_j(M_j) \tag{3}$$

We assume that no additional benefit can be gained by repeatedly monitoring the same traffic. Thus we can express M_j in either of two ways:

$$M_j = 1 - \prod_i (1 - m_{ij}) \tag{4}$$

$$M_j = \sum_i m_{ij} \tag{5}$$

Equation (4) models a monitor that independently samples packets. Equation (5) models monitors that mark sampled packets and only sample unmarked packets. We assume that $u_j()$ is non-decreasing concave. A simple linear function may be a reasonable candidate. Alternatively, we might choose a form of $u_j()$ that accounts for sampling errors, in which case $u_j()$ will also be a strictly non-decreasing concave function [9, 10].

III. A SET OF PASSIVE MONITORING PROBLEMS WITHOUT SAMPLING

In this section, we introduce several monitoring problems under the assumption that each monitor collects *all* packets of monitored flows, i.e., $m_{ij} = 1$ or 0 for all i, j. We address 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 problems and that there exist approximate algorithms that give results close to optimum. Table II summarizes the set of passive monitoring problems described in this section.

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	minimum SC
MDOCP	variable	variable	no	min placement+operating cost	uncapacitated FLP

TABLE II
A SET OF PASSIVE FLOW MONITORING PROBLEMS WITHOUT SAMPLING

A. Budget Constrained Maximum coverage problem without sampling (BCMCP)

The objective of BCMCP is to maximize the monitoring reward without violating a constraint on the total deployment cost. Each monitor has an associated deployment cost and there is a limited budget to cover this cost. Initially, 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 by optimally assigning monitors to links.

Let B denote the maximum budget for the total deployment cost, and variable x_j indicate whether flow j is being 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:

$$
\begin{array}{lll}
\text{Maximize} & \sum_{j \in D} u_j(x_j) & \\
\text{subject to} & x_j & \leq \displaystyle\sum_{i: j \in S_i} y_i, \;\; j \in D \\
& \sum_{i \in L} f_i y_i & \leq B \\
& y_i & \in \{0,1\}, \;\; i \in L \\
& x_j & \in \{0,1\}, \;\; j \in D
\end{array}
$$

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 $S = \{S_1, S_2, \ldots, S_m\}$ 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 $S' \subseteq S$, such that the total weight of elements covered by S' is maximized, while the total cost of elements in S' is less than a given budget L [11]. The reduction can be done in the following way. S maps to the set of links; the associated cost c_i is mapped to the deployment cost f_i; and S_i maps to the set of flows 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 $u_j(x_j)$ of each flow j and the budget L is mapped to the budget constraint for the total deployment cost, B.

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

We propose a $(1-1/e)$-approximation algorithm for BCMCP, which is adapted from greedy heuristic for the budgeted MCP proposed by S. Khuller et al. [11]. For a special case of the problem, where each subset has unit cost, the simple greedy heuristic picks, at each step, a subset, S_i, maximizing the utility value of the uncovered flows. For the general case in which each subset 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. The second candidate solution is obtained using simple greedy heuristics with a different seed element. More specifically, for each subset $G \subseteq S$ of cardinality k that has cost at most B, G is included as a very first element to a candidate solution. After that, the candidate solution is augmented with other subsets in S, which greedily maximize the utility value of the uncovered 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 utility value of the flows covered by set S_i, but not covered by any set in G. Note that the parameter k is a tunable parameter that is chosen by a user. 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 [11].

```
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 $S_i \in U$ 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

B. 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 given a monitoring reward requirement. 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 utilities of each flow are set equal to a constant value, MDCP simply minimizes the deployment cost while the number of flows being monitored is equal or greater than the given monitoring reward, K. If the utility of each flow represents the traffic demand of the flow, MDCP is defined to minimize the deployment cost subject to the constraint that some minimum fraction of packets be monitored.

The integer linear program formulation of MDCP is as follows. We want to find an assignment for the variables y_i, that:

$$\begin{aligned} \text{Minimize} \quad & \textstyle\sum_{i\in L} f_i y_i, \\ \text{subject to} \quad & \\ x_j \quad & \leq \sum_{i:j\in S_i} y_i, \;\; j \in D \\ \textstyle\sum_{j\in D} u_j(x_j) \quad & \geq K \\ y_i \quad & \in \{0,1\}, \;\; i \in L \\ x_j \quad & \in \{0,1\}, \;\; j \in D \end{aligned}$$

Since MDCP is the dual problem of BCMCP, MDCP also is NP-hard. It can also be directly shown that MDCP is NP-hard by constructing a dependency matrix whose rows represent links and whose columns represent network flows. Here, we only prove the case that the utilities of all monitored flows are equal; the NP-hardness of the other cases with different utilities can be proven in the similar way. Each entry (i, j) in the matrix can take binary values $\{0, 1\}$ such that if flow j traverses link i, entry (i, j) has value 1 and has value 0 otherwise. We need to choose a subset of rows that covers at least K columns with the total cost of selected subset minimized. This is exactly a weighted version of the partial set cover problem, which is known to be NP-hard.

For the general case, we can get an $(log(|D|)+1)$-approximation solution by adapting the approximation algorithm for partial K-set cover problem proposed in [12]. 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 are integer values. For each $e_j \in S_i$, if $u_j(e_j) = M$ then we replace e_j in S_i with M flows $e_{j1}, e_{j2}, \ldots$, and e_{jM} where $u(e_{jm}) = 1$, $m = 1, \ldots, M$. For example, take $S_i = \{e_1, e_2, e_3\}$, where $u(e_1) = 3, u(e_2) = 2$, and $u(e_3) = 1$. Then, we replace the original elements $\{e_1, e_2, e_3\}$ with new elements with unit utility values such that $S_i = \{e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{31}\}$. 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 in order to obtain a K-cover and the number of uncovered flows in each subset S_i.

```
Approximation algorithm ($(K)$)
(0)  Replace each flow $j \in S_i$, where $u_j(x_j) = m$ for $x_j = 1$, 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

C. 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 station is deployed. Unlike BCMCP and MDCP, each monitor is allowed to *selectively* monitor a set of flows, instead of monitoring all the flows. If a flow is monitored, however, all packets in that flow will be 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 objective function is minimized:

$$\begin{array}{lll} \text{Minimize} & \sum_{i\in L} f_i y_i & + \sum\limits_{i\in L} y_i c_i \sum\limits_{j\in D} \rho_j x_{ij} \\ \text{subject to} & & \\ & x_{ij} & \leq y_i,\ \ i \in L, j \in S_i \\ & \sum_{i\in L} x_{ij} & = 1,\ \ j \in D \\ & x_{ij} & \in \{0,1\},\ \ i \in L, j \in S_i \\ & y_i & \in \{0,1\},\ \ i \in L \end{array}$$

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 distances between them denoted as c_{ij}, i, j=$1, \ldots, n$. We may open a facility at potential facility locations, $F \subseteq N$; building a facility at location $i \in F$ has an associated non-negative cost f_i. We also have 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. We assume that these costs are 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 that is the sum of facility opening cost and the total shipping cost [13, 14]. 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 [14].

We can map our problem to the uncapacitated facility location problem in the following way. Let's suppose that we have a set of flows M and a set of link 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, $c_{ij} = \infty$.

IV. 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. The sampling rate for each flow at each monitoring station can be adjusted independently. 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 sampling version of the BCMCP problem considered in section III.

The objective of BCMCP-S is to maximize the total utility of fractional flows being sampled without violating the budget constraint for the monitors' deployment and operating cost. More specifically, we have limited budgets to cover the deployment cost and operating cost. Once a monitor is deployed on a link, a subset of the flows carried by the link is sampled by the monitor. The sampling rate per flow at each monitor is controlled independently. In this setting, 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}, that:

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

In general, it is difficult to directly obtain an optimal solution, because the y_i's are integer variables and the $u_j(M_j)$ may be nonlinear functions [15]. For example, if we assume that monitors sample flows independently of each other, the objective function can be represented as $\sum_{j\in D}(u_j(1-\prod_{i\in L}(1-m_{ij})))$. To obtain an optimal solution, we can apply the branch-and-bound algorithm combined with gradient projection method to reduce the effect of combinatorial explosion [15]. However, we still need a faster method to compute solutions since the branch-and-bound method cannot guarantee a computation time that is polynomial in the number of monitors. Here, we obtain an approximate solution as follows. We solve BCMCP-S approximately via a two-stage algorithm consisting of a greedy algorithm for integer variables and a gradient projection method for 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 the 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 */
$H_1 \leftarrow$ argmax $\{w(G)$ , such that $G \subseteq S$, $|G| \leq k$, and $c(G) \leq B_1\}$
$H_2 \leftarrow 0$
For all $G \subseteq S$, such that $|G| = k$ and $c(G) \leq B_1$ do
    $U \leftarrow S \backslash G$
    Repeat
        Select $S_i \in U$ that maximizes $\sum_{j \in S_i} (u_j(1)/\rho_j)/c_i$
        If $c(G) + c_i \leq B_1$ 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 $w(H_1) > w(H_2)$, then $H \leftarrow H_1$, otherwise, $H \leftarrow H_2$
For all $S_i$, if $S_i \in H$, then $y_i \leftarrow 1$, otherwise, $y_i \leftarrow 0$
    /* Stage 2 */
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

V. 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 in comparison with the optimal solutions. We also investigate the problem of determining the number of monitors needed to achieve a specific level of monitoring reward. We first describe the specific parameter settings for the two problems and then show the results of our evaluation.

A. Simulation parameter settings

1) Network topology, traffic matrix, and routing settings: We use both synthetic network topologies and a real ISP topology for our study. More specifically, we use the PoP-level topology of Cable&Wireless, as inferred by

[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 [16].

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

TABLE III
PARAMETER SETTINGS

the Rocketfuel project [17, 18]. Our synthetic topologies, consist of random topologies and Transit-Stub topologies generated by GT-ITM [19]. 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 [20].

In [20], 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 lead to 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 and so the demands are modeled as follows. For each x, two random numbers are chosen: O_x, $D_y \in [0, 1]$. 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 [20].

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 routing path of each flow is determined by a shortest path routing algorithm assuming that a single path is always taken by each flow.

2) Utility functions and cost assignment: Table III lists the utility function $u_j()$ for each flow j, the deployment cost f_i for each monitor at link i, and the values of all of the parameters. In Table III, the traffic demand of flow j, $\rho_j x_j$, is taken as the utility $u_j(x_j)$ for BCMCP. We take the fraction of monitored packets in flow j, M_j, as the utility $u_j(M_j)$, which we call 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 unit value 1 for deployment costs, f_j.

B. 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 a random network of 10 routers listed in Table IV. The x-axis represents the number of deployed monitors and the y-axis the fraction of packets monitored. Though BCMCP is an NP-hard problem, we can compute optimal solutions for small problem sizes and compare them to solutions produced by our heuristics. The graph in Figure 4 shows that the optimal solutions achieve higher reward than the greedy solutions, although

Type	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

TABLE IV
GT-ITM TOPOLOGIES AND THE TOTAL NUMBER OF FLOWS

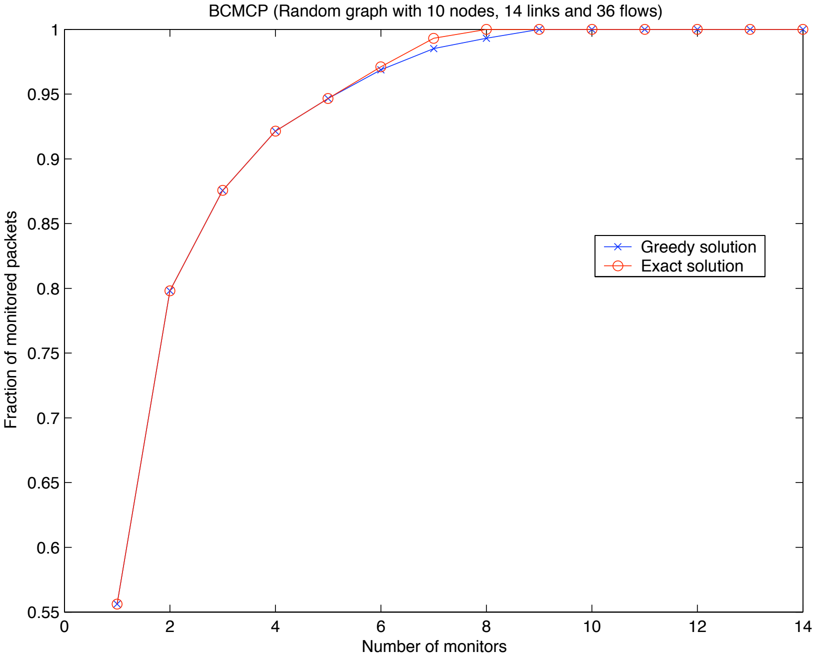

Fig. 4. BCMCP in a 10 node, 14 link, and 36 flow GT-ITM Random topology

the differences between the optimal and heuristic solutions are quite small. In this graph, we 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 IV. Because of the combinatorial explosion, we could only compute optimal solutions for up to 5 monitors for this problem. Interestingly, the greedy solution for 3 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 III, we have shown that $(1 - 1/e)$ is the theoretical bound of approximation ratio of the greedy solution to the optimal solution in BCMCP. However, in Figure 4 and 5, we observe that the greedy algorithm comes much closer to optimal. We conjecture that the better approximation performance is made possible by various factors such as shortest-path routing, the hierarchical structure of the topologies, and all-pair traffic demands among nodes. One more 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 cover

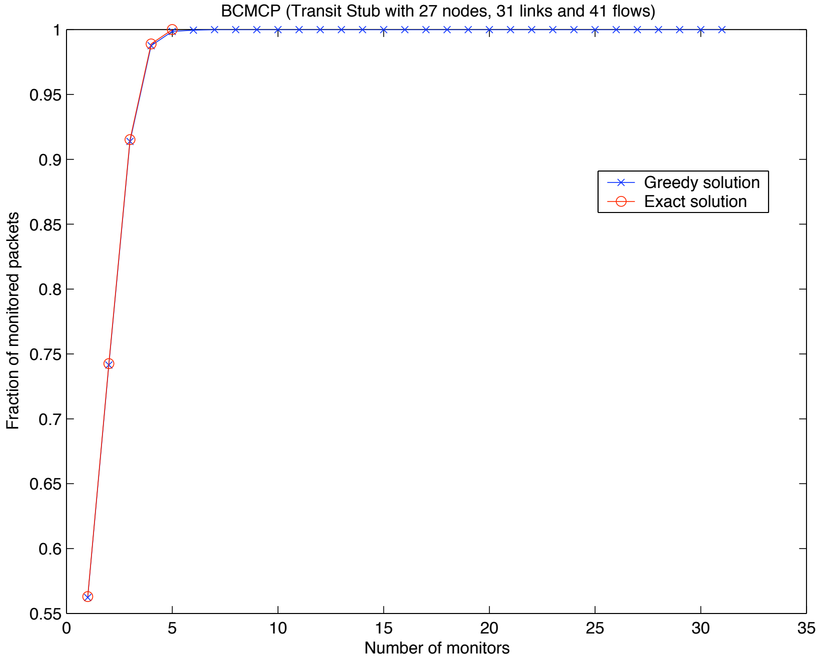

Fig. 5. BCMCP in a 27 node, 31 link, and 41 flow GT-ITM Transit-Stub topology

90% of the packets, only 10% of the locations require monitoring in latter case, while, about 29% of the locations require monitoring in the former case. We conjecture that this is due to the fact that 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 graph of 100 nodes. Again we could not compute exact solutions for all possible number of monitors. This graph also shows the decreasing monitoring reward gain. In addition, in Figure 6 a smaller fraction of monitoring locations is required by the network than the 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 use the PoP-level topology [4] of Cable & Wireless as inferred by the Rocketfuel project [17, 18]. 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 [20]. Also, we use a shortest path routing algorithm to find the routing path for each flow, assuming that a single path is always used. In Figure 7, we also observe similar trends such as a diminishing reward gain.

C. Simulation results for BCMCP-S

Figure 8 shows simulation results for the BCMCP-S approximation applied to the 10-node random graph listed in Table IV. In addition to the parameter settings in Table III, we select a budget, B_2, for the total operating cost such that it is impossible to sample 100% of packets of the flows covered by monitors for any set of monitors that satisfy the budget constraint, B_1. 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

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

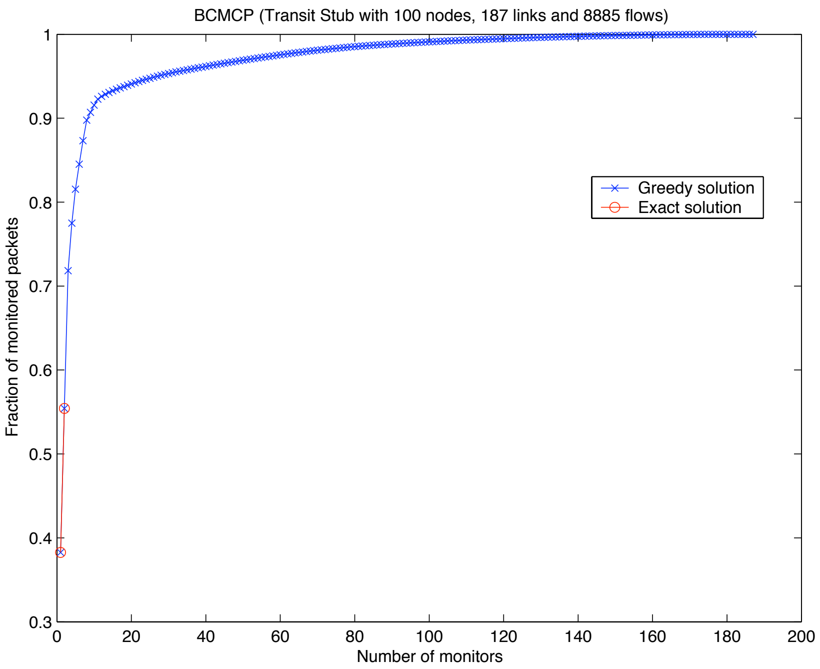

Fig. 6. BCMCP in a 100 node, 187 link, and 8885 flow GT-ITM Transit-Stub topology

set of monitors. The graph shows that the greedy solution is quite close to the optimal solution. In Figure 9, we use the same random topology as in Figure 8, but replace the linear utility with an exponential utility function. Specifically, the objective function is $u_j(M_j) = (1 - exp(-5.0 * M_j))$, where $M_j = (1 - \prod_i (1 - m_{ij}))$. We again observe that the greedy solution is close to the optimal solution. In addition, both Figure 8 and 9 show 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 latter case that uses the exponential function more flows are monitored (although smaller sampling rates are obtained on average) than in the former case that uses the linear function.

VI. Related Work

Several recent efforts have addressed the placement problem of *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 those areas. We summarize prior work in this section.

A. Number and location of tracers in the Internet

To obtain topology and performance measurement such as link delay and existence of faults etc. measurement points that send “active probe” messages may be used. The location of these measurement devices or beacons has been determined according to various heuristics in the literature [2–5]. 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. [6] showed the minimal number of required beacons on a network and their near-optimal locations. Barford

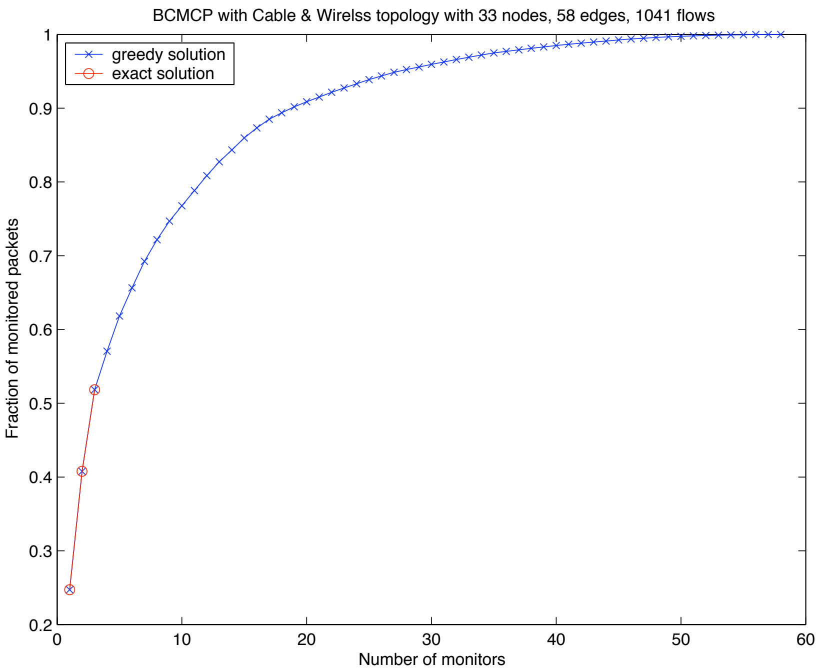

Fig. 7. BCMCP in Cable&Wireless PoP-level topology from Rocketfuel dataset

et al. [21] presented empirical observation that the marginal utility of adding additional active measurement sites declines rapidly after the second or third site. In other words, both works show that a relatively small number of active measurement points is generally sufficient to obtain an accurate network topology [6, 21]. This conclusion is similar to our observation that only a small number of passive monitors is necessary to achieve high monitoring coverage. However, these earlier works were concerned about active, rather than passive, monitors.

B. 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 an approximation to the problem of placing route-based packet filters on routers to prevent distributed DoS Attack. Since the vertex cover problem is known to be NP-hard, they investigated several greedy heuristics. Also, they 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 enough 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.

C. Network tomography

The objective of Minimum cost Multicast Tree Cover Problem (MMTCP) in [23] 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

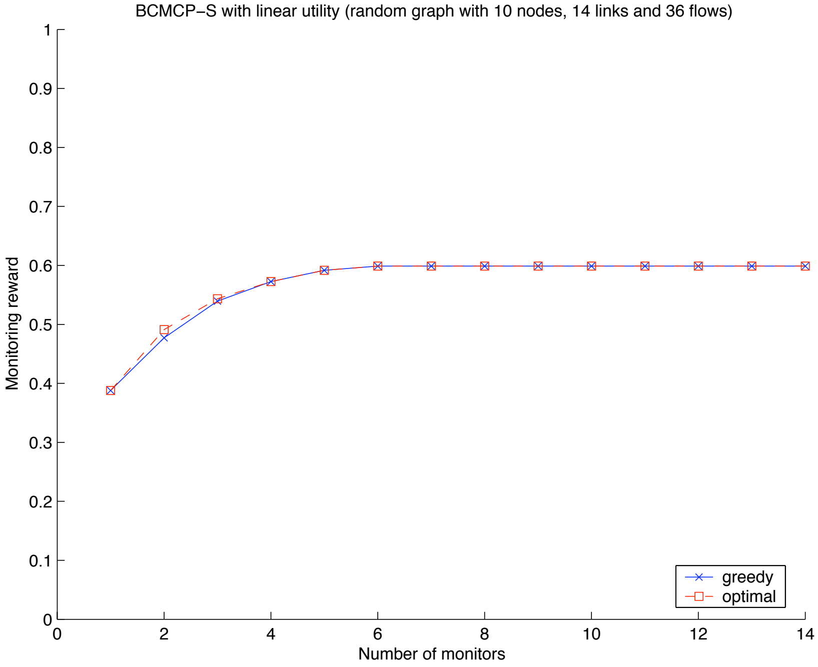

Fig. 8. BCMCP-S in a 10 node, 14 link, and 36 flow GT-ITM Random topology

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 [23].

D. Sampling strategies

Sampling methodologies at a single measurement point [1, 9, 24] 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 $u_j(M_j)$ 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 [9, 24], 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 [1], 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 $p_z(x) = min\{1, x/z\}$ [5] minimizes the cost function $C_z(p)$.

VII. Conclusion

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

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

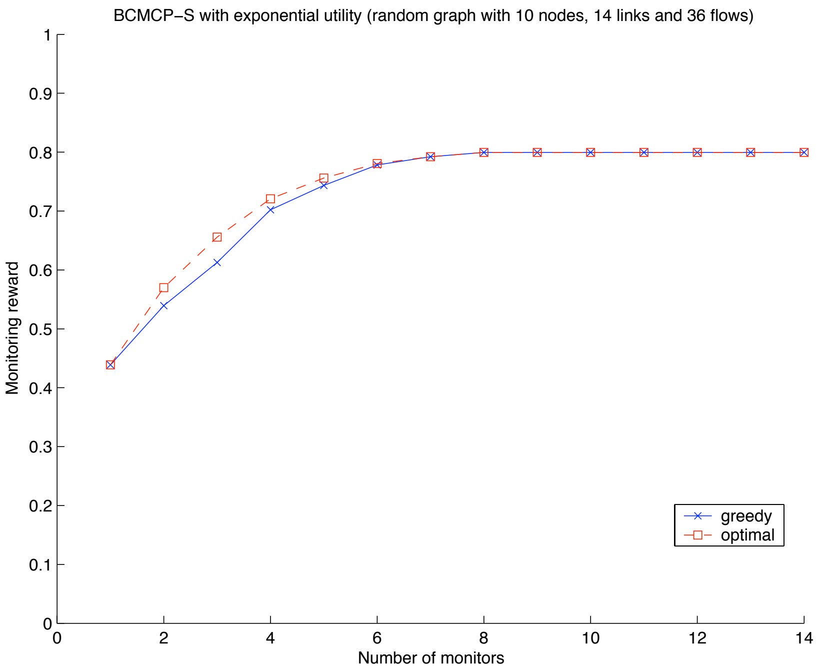

Fig. 9. BCMCP-S in a 10 node, 14 link, and 36 flow GT-ITM Random topology

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 and coverage 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 topology. 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. Also, we are considering the case of route changes caused by link failures. 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.

References

[1] N. Duffield, C. Lund, and M. Thorup, "Learn more, sample less: control of volume and variance in network measurement." 2002.

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

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

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

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

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

[7] 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.

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

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

[10] 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.

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

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

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

[14] F. Chudak and D. Shmoys, "Improved approximation algorithms for the unpacacitated facility location problem," *ACM SIAM Journal on computing*, no. 1, 2003.

[15] 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.

[16] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

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

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

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

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

[21] P. Barford, A. Bestavros, J. Byers, and M. Crovella, "On the marginal utility of network topology measurements," in *Proc. ACM SIGCOMM Internet measurement workshop*, 2001.

[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. 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.

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