Joint Optimization of Monitor Location and Network Anomaly Detection

Emna Salhi, Samer Lahoud, Bernard Cousin

Joint Optimization of Monitor Location and Network Anomaly Detection

Emna Salhi, Samer Lahoud, Bernard Cousin
IRISA / University of Rennes 1, France
{emna.salhi, samer.lahoud, bernard.cousin}@irisa.fr

***Abstract*—Achieving cost-effective systems for network performance monitoring has been the subject of many research works over the last few years. Most of them adopt a two-step approach. The first step assigns optimal locations to monitors, whereas the second step selects a minimal set of paths to be monitored. However, such an approach does not consider the trade-off between the optimization objectives of each step, and hence may lead to sub-optimal usage of network resources and biased measurements.**

In this paper, we propose to evaluate and reduce this trade-off. Toward this end, we come up with two ILP formulations for a novel monitoring cost model that apply for both passive and active monitoring. The aim is to minimize the monitor location cost and the anomaly detection cost jointly, thereby obtaining a monitoring solution that minimizes the total monitoring cost. Simulation results illustrate the interplay between the optimization objectives and evaluate the quality of the obtained monitoring solution.

I. Introduction

Monitoring cost includes a monitor deployment cost and an operational cost. The monitor deployment cost expresses the effective cost of deploying hardware and software monitoring devices. The operational cost quantifies the overhead on the underlying network due to communications between monitors and the *Network Operations Center* (NOC). It also quantifies, for active monitoring, the burden on links generated by the injected monitoring flows. Most existing works on network monitoring adopted a two-step scheme: the first step, known as monitor location step, aims at minimizing the monitor deployment cost; whereas the second step, known as path selection step, aims at minimizing the operational cost.

A trivial optimization of the first step consists in reducing the number of monitors. Several works proposed schemes to place as few monitors as possible at strategic locations such that all links are covered (*e.g.* [5]). Works in [1]-[4] addressed the optimization of path selection step. Given an optimal set of monitor locations, they proposed inference schemes that monitor a small set of paths toward minimizing the communication cost. One of the most common approaches is to perform the monitoring task over two phases: anomaly detection phase and anomaly localization phase (*e.g.* [1]-[2]). The goal is to reduce monitoring overhead when network behaves well during the detection phase by monitoring few paths. All these works decouple the monitor location problem from the path selection problem, and hence do not consider the impact of the number and the locations of monitors on the quality of monitored paths.

Recently, Zhao et al. [1] argued that link and monitor capacities to handle monitoring flows should be considered while selecting monitor locations. The authors claimed that the problem is quite complex; and proposed a multi-round monitoring scheme that reduces the complexity by a factor of the number of rounds. The major limitation of such an approach is that it increases the delay to detect anomalies by a factor of the number of rounds. In this paper we investigate and reduce the trade-off between the optimization objectives of the two steps. Toward this end, we propose two different ILP formulations that model a joint optimization of monitor location and network anomaly detection problems. Given a set of operational constraints, our ILPs provide optimal locations for monitors and optimal set of paths to be monitored that minimize the total monitoring cost and satisfies the constraints. The two ILPs were solved on randomly generated network topologies, in order to investigate the complexity of the problem and to obtain a deeper understanding of the interpaly between the optimization objectives and their impact on the quality of the solution.

II. Problem Formulation

We model the network as an undirected graph $G = (N, E)$, where N denotes the set of nodes, and E denotes the set of bidirectional edges that represent the set of links connecting nodes. We denote by P the set of non-looping network paths. A solution for network performance monitoring consists of two parts: a set of locations where to deploy monitors, and a set of paths that are to be monitored to detect and localize anomalies. In this paper, we are not interested on the localization of anomalies. We adopt the most common approach of anomaly detection that is monitoring a covering path set that do not distinguish link anomalies (*e.g.* [1]-[2]). We consider a centralized monitoring infrastructure where the NOC, which has a global view of the network topology, ensures the monitor location and path selection tasks. A monitored path is defined to be a sequence of links carrying monitoring flows. We define the monitor location cost and the anomaly detection cost as follows:

-**Monitor location cost:** Let Cd be the cost of deploying a monitor in the network and Y_n a binary indicator if a

This work has been supported by Alcatel-Lucent Labs France, under the grant n. 09CT310-01

monitoring device is located on node n , the total monitor location cost can be expressed as follows:

$$Cd \sum_{n \in N} Y_n \tag{1}$$

-**Anomaly detection cost:** it includes two costs, a communication cost and a link measurement cost. The communication cost is the cost associated with the communications between monitors and the NOC, *e.g.* to synchronize monitors, ship measurements. Toward minimizing this cost, monitors should be located as near as possible to the NOC. Let D_n be the distance in number of hops of node n to the NOC, the total communication cost is:

$$\sum_{n \in N} D_n Y_n \tag{2}$$

The link measurement cost expresses the burden on network links due to the injected monitoring flows. This cost is zero for passive monitoring. Let Cl_l be the cost of injecting a monitoring flow along link l and R_l an integer counter that indicates the number of monitoring flows traversing l. Cl_l is proportional to the load of link l. The aim is to avoid redundant measurements of overloaded links. the link measurement cost can be expressed as follows:

$$\sum_{l \in E} Cl_l R_l \tag{3}$$

We provide an example to illustrate the trade-off between the optimizations of these costs. We consider two monitoring scenarios run on a small network depicted in Fig.1 and Fig.2. In each scenario, we set the number and positions of monitors, and then we compute an optimal set of paths to be monitored. the chosen path set must cover all network links while minimizing redundant measurements. In the first scenario (Fig.1), we locate two monitors on nodes 2 and 8. A solution that matches this setting is S1 = {P1, P2)}. In the second scenario (Fig.1), we locate three monitors on nodes 1, 6 and 8. S2 = {P3, P4, P5} is an optimal solution.

In the first scenario, 3 links are monitored twice; they belong each to two monitored paths. The deployment of an additional monitor in the second scenario reduces the number of bi-monitored links to one link. We conclude that the less monitors are deployed, the more redundant measurements of links we obtain. In the sequel, we introduce two ILP formulations that minimize jointly the costs given by (1), (2) and (3). The first formulation is a path based ILP that takes as input the set of candidate paths. The second formulation is a link-flow based ILP that avoids the pre-computation of the set of network paths.

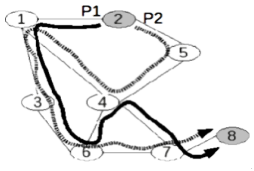


Fig. 1. Scenario 1: Two monitors are deployed

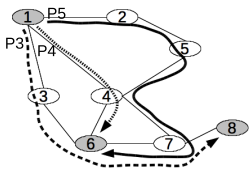


Fig. 2. Scenario 2: Three monitors are deployed

A. Path based ILP Formulation

Let us denote by Cm_n the sum of Cd and D_n. Our path based ILP formulation aims at minimizing the total monitoring cost given by the sum of (1), (2) and (3):

$$\textit{Minimize: } \alpha \sum_{l \in E, p \in P} Cl_l \delta_{lp} Z_p + \beta \sum_{n \in N} Cm_n Y_n \tag{4}$$

Z_p is a binary variable that indicates if path p is monitored, and δ_{lp} is a binary constant parameter that indicates if path p traverses link l. The number of monitoring flows traversing link l is given by the sum $\sum_{p \in P} \delta_{lp} Z_p$. α and β are positive weights that determine the relative importance of the optimization components of the above cost function.

The objective function is subject to the following constraints:

$$\sum_{p \in P} \delta_{ep} Z_p \geq 1; \quad \forall e \in E \tag{5}$$

$$Y_n \geq \delta_{np} Zp; \quad \forall n \in N, \forall p \in P \tag{6}$$

δ_{np} is a constant binary parameter that indicates if node n is an end node of path p. Constraints (5) guarantee that each network link belongs to some monitored path, whereas constraints (6) ensure that the end nodes of each monitored path are selected as monitors.

We can show that this ILP formulation is NP-hard by mapping it to the uncapacitated facility location problem. However, we do not provide a demonstration due to lack of space.

B. Link-flow based ILP formulation

We expect that the path based ILP would not scale to large networks where the number of paths is drastically high. In an attempt to overcome this limitation, we propose a link-flow based ILP formulation that avoids the pre-computation of the set of network paths. Beside the basic monitoring constraints, *i.e.* covering the network links and selecting the end nodes of paths carrying monitoring flows as monitors, we formulate constraints that avoid forming looping paths and ensure flow conservation at nodes. We use interchangeably the terms *path* and *flow* to designate a path that is candidate to carry monitoing flows. Let $A = \{(i \rightarrow j), (j \rightarrow i); \forall (i,j) \in E\}$ be a virtual arc set, and let $Cl_{(i \rightarrow j)}$ denotes the cost of monitoring arc $(i \rightarrow j)$. We have $Cl_{(i \rightarrow j)} = Cl_{(j \rightarrow i)} = Cl_{(i,j)}$. The flows are modeled using a set of binary variables $\{X_{i \rightarrow j}(n,n'); (i \rightarrow j) \in A, (n,n') \in N^2\}$, each variable $X_{i \rightarrow j}(n,n')$ expresses whether the flow travelling between the pair of nodes (n,n') and crossing the arc $(i \rightarrow j)$ is monitored. The link-fow based ILP reads as follows:

$$\textit{Minimize: } \alpha \sum_{(i,j) \in E, (n,n') \in N^2} Cl_{(i,j)} [X_{i \rightarrow j}(n,n') + X_{j \rightarrow i}(n,n')] + \beta \sum_{n \in N} Cm_n Y_n \tag{7}$$

Subject to the following constraints:

1) Each network link must be monitored at least once:

$$\sum_{(n,n') \in N^2} X_{i \rightarrow j}(n,n') + X_{j \rightarrow i}(n,n') \geq 1; \quad \forall (i,j) \in E \tag{8}$$

2) Multiple monitoring flows might be carried between a pair of nodes. We define a set of integer variables $\{W_{(n,n')}; (n,n') \in N^2\}$ to quantify the number of monitoring flows travelling between each pair of nodes. Let $IN(v)$ and $OUT(v)$ be the set of arcs entring node v and the set of arcs leaving node v, respectively. The flow conservation constraints are, hence, expressed as follows:

$$\sum_{i\to j \in OUT(v)} X_{i\to j}(n,n') - \sum_{i\to j \in IN(v)} X_{i\to j}(n,n') = \begin{cases} W_{(n,n')} & \text{iff } v = n \\ -W_{(n,n')} & \text{iff } v = n' \\ 0 & otherwise \end{cases} ; \quad \forall v,n,n' \in N \quad (9)$$

3) The following constraints ensure that the end nodes of paths carrying monitoring flows are selected as monitors:

$$Y_n \geq W_{(n,l)} + W_{(l,n)}; \quad \forall n \in N, \forall l \in E \quad (10)$$

4) Toward preventing looping flows, we define a set of integer variables $\{H_{(n,n')}(i); n,n', i \in N\}$. $H_{(n,n')}(i)$ specifies the number of hops separating node i visited by a flow travelling between the pair of nodes (n,n') from its originating node n. The idea is to force the flows to travel through nodes in an ascending order of the values of their hop variables, which prevents them from looping. We formulate the looping constraint as follows:

$$H_{(n,n')}(n) = 0; \quad \forall (n,n') \in N^2 \quad (11)$$

$$\begin{aligned} 1 - X_{i\to j}(n,n') + \frac{H_{(n,n')}(j) - 1 - H_{(n,n')}(i)}{K} \geq 0 \\ 1 - X_{j\to i}(n,n') + \frac{H_{(n,n')}(i) - 1 - H_{(n,n')}(j)}{K} \geq 0 \end{aligned} ; \quad \forall (i,j) \in E, (n,n') \in N^2 \quad (12)$$

$$H_{(n,n')}(n') \leq |N| - 1; \quad \forall (n,n') \in N^2 \quad (13)$$

III. Evaluation

In this section, we present our evaluation methodology, metrics, and simulation results.

A. Methodology and Metrics

We evaluated our ILPs using Cplex11.2 [7] running on a PC equipped with an Intel(R) Core(TM)2 Duo processor and 3.9 GB of RAM. All results are the mean over 20 simulations on random topologies generated using the topology generator BRITE (AS level, Waxman model) [6]. Table I depicts a summary of the main characteristics of topologies considered in our evaluation. We devised and implemented an algorithm that computes the set of paths of an input topology. As we have anticipated owing to the complexity of the problem, we failed to compute the path set for TOP(12, 41) due to memory failure. We considered an active monitoring scenario where all the network paths are candidate to be monitored and all the nodes are candidate to hold monitors, and we assumed that nodes are equidistant from the NOC. The values of Cl_l and Cm_n are set to 1 $\forall l \in E$ and $\forall n \in N$, respectively. We considered the following metrics for the evaluation of the ILPs:

- Gap-to-optimality: it expresses the gap between the obtained solution and the optimal solution estimated by the solver. We chose to present this metric instead of the value of the objective function, because for large topologies, the solver failed to compute an optimal solution within a reasonable time. This metric allowed us to compare the performance of the two ILPs and to validate our expectations; *(i) The path based ILP is quite greedy for memory, because it must manage the network path set given as input, (ii) The link-flow based ILP requires high processing capacity to handle the huge number of variables and constraints*.
- Toward studying the trade-off between minimizing the monitor cost and minimizing redundant measurements of links; we tuned the values of $\frac{\beta}{\alpha}$, and investigated the quality of the obtained solutions. We considered three settings: $\frac{\beta}{\alpha} = 1$; $\frac{\beta}{\alpha} = 10^3$, *i.e.* $\beta >> \alpha$; and $\frac{\beta}{\alpha} = 10^{-3}$, *i.e.* $\beta << \alpha$. For each setting, we have investigated the following metrics: number of deployed monitors, number and average length of monitored paths, and number of redundant measurements of links.

TABLE I
Summary of the topologies considered in the evaluation

Topology	# of nodes	# of links	# of paths
TOP(6, 10)	6	10	162.5
TOP(8, 18)	8	18	3176.9
TOP(10, 31)	10	31	209235.2
TOP(12, 41)	12	41	*

B. Results

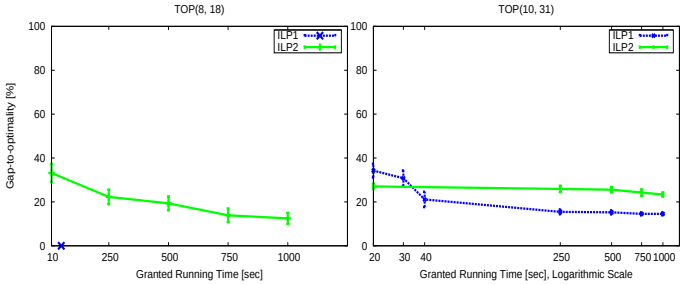
In the sequel, we refer to the path based ILP as ILP1 and the link-flow based ILP as ILP2.

1) Evaluation of the performance of the ILP formulations: In this section, we present results for $\alpha = \beta$. Tab. III-B1 presents the gap-to-optimality (GTO) and the CPU running times (RT) for the smallest topologies, *i.e*. TOP(6, 10), and the largest topologies, *i.e.*TOP(12, 41). We notice that for TOP(6, 10), the two ILPs generated optimal solutions (GTO = 0%). However, the running times show that the resolution of ILP1 is much easier than the resolution of ILP2. This validates our assertion that ILP2 is more demanding in processing capacities. This observation is confirmed in Fig.3(a), which plots the GTO versus the granted RT for TOP(8, 18). Indeed, this figure shows that ILP1 was able to obtain an optimal solution in 50.82 seconds, while after 10^3 seconds, ILP2 provided a solution with nonzero GTO. Fig. 3(b) plots the GTO versus the granted RT for TOP(10, 31). It shows that the two ILPs failed to generate optimal solutions within 10^3 seconds. We observe that when the granted RT is small, the solutions provided by ILP1 are worse than those provided by ILP2; while when the granted RT is large enough, ILP1 performs better than ILP2. This is possibly due to the large number of paths. Compared to the results obtained by ILP1 for TOP(8, 10), we notice that the GTO of those obtained for TOP(10, 31) goes up dramatically. This explicitly verifies that ILP1 is quite sensitive to the network size. The results for TOP(12, 41) further validates this observation. Indeed, Tab. III-B1 shows that ILP1 failed to provide a feasible solution due to memory failure, whereas ILP2 generated a solution only 25.01% worse than the optimal within 10^3 seconds.

TABLE II

EVALUATION RESULTS FOR TOP(6,10) AND TOP(12,41). GTO DENOTES GAP-TO-OPTIMALITY, RT DENOTES CPU RUNNING TIME

	ILP1		ILP2	
Topology	**GTO[%]**	**RT[sec]**	**GTO[%]**	**RT[sec]**
TOP(6, 10)	0	0.03	0	20.5
TOP(12, 41)	Out of Memory		25.01	1000



(a) (b)

Fig. 3. Gap-to-optimality Vs. Granted Running Time

2) Evaluation of the trade-off between the cost optimization components: now we investigate the quality of the obtained solutions versus the ratio $\frac{\beta}{\alpha}$ for TOP(6, 10), TOP(8,18), and TOP(10,31). As ILP2 failed to generate optimal solutions within a reasonable time for TOP(8, 10) and greater, we limit our simulations on ILP1. For TOP(10, 31), we show results obtained within 10^3 seconds. Fig.4(a) plots the average number of monitors versus network topology and weight ratio. As expected, the figure shows that when $\beta >> \alpha$, only two beacons are deployed for all topologies. This is the minimal number of monitors required to monitor a path. Obviously, the monitored paths, which have the same end nodes, are likely to overlap. This is verified in Tab.III-B2, which shows that the percentage of redundant measurements of links ranges from 6% to 10%. Fig.4(a) shows that the number of monitors deployed when $\beta << \alpha$ is larger by several orders than those deployed when $\beta >> \alpha$, however, it is lower than to the total number of nodes. This is because, the number of monitors is also minimized in a way that minimizes the total monitoring cost. Clearly, the additional monitors are deployed to remove path overlaps. Tab.III-B2 validates this assertation. Indeed, it shows that 100% of network links are monitored once for TOP(6,10) and TOP(8,18), and only 3.55% of network links are monitored twice for TOP(10,31).

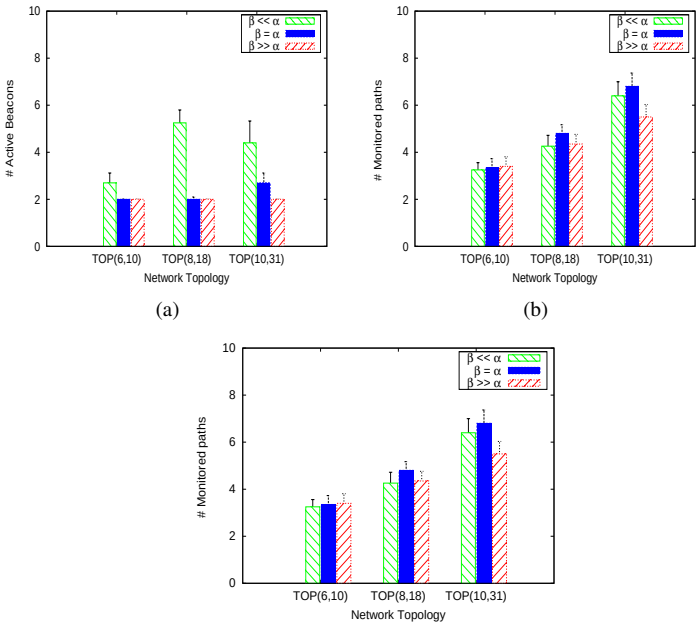
The above analysis results suggest that there is a trade-off between minimizing the number of monitors and minimizing redundant measurements of links. However, the joint optimization of these two objectives succeeds to reduce the trade-off. Indeed, Tab.III-B2 shows that less than 10% of links are monitored twice when $\beta >> \alpha$, and Fig.4(a) shows that only 60% of nodes are selected as monitors when $\beta << \alpha$.

Surprisingly, Fig.4(c) and Fig.4(b) show that the average number and the average length of monitored paths are barely sensitive to the value of the weight ratio. This meets our observation that considering the number and the length of monitored paths as the only criteria for path selection does not necessarily lead to an optimal monitoring solution.

TABLE III

PERCENTAGE OF REDUNDANT MEASUREMENTS OF LINKS. % SM (SINGLE MONITORING) DENOTES THE % OF LINKS MONITORED ONCE, % DM (DOUBLE MONITORING) DENOTES THE % OF LINKS MONITORED TWICE

	$\beta << \alpha$		$\beta = \alpha$		$\beta >> \alpha$	
	% SM	% DM	% SM	% DM	% SM	% DM
TOP(6,10)	100	0	94	6	94	6
TOP(8,18)	100	0	90.56	9.44	90	10
TOP(10,31)	96.45	3.55	91.61	8.39	92.34	7.66



(a) (b)

(c)

Fig. 4. Quality of Monitoring Solutions Vs. Weight Ratio and Network Topology

IV. CONCLUSION

In this paper we advocate a monitoring cost model that reduces the trade-off between minimizing the monitor location cost and minimizing the anomaly detection cost. We introduce a path based ILP formulation and a link-flow based ILP formulation, each optimizes jointly the two costs. Results show that the path based ILP is quite greedy for memory, and the link-flow based ILP is quite greedy for CPU. Hence, the two ILPs could not be used to compute monitoring solutions for large networks. However, we succeeded to validate our observations on small networks. One goal of our future work is to devise heuristics for our optimization model.

REFERENCES

[1] Y. Zhao, Z. Zhu, Y. Chen, D. Pei, and J. Wang, "Towards efficient large-scale VPN monitoring and diagnosis under operational constraints", IEEE INFOCOM, 2009.

[2] P. Baford, N. Duffield, A. Ron, and J. Sommers, "Network performance anomaly detection and localization", IEEE INFOCOM, 2009.

[3] K.V.M Naidu, D. Panigrahi, and R. Rastogi, "Detecting anomalies using end-to-end path measurements", IEEE INFOCOM, 2008.

[4] S. Argawal, K.V.M. Naidu, and R. Rastogi, "Diagnosing link-level anomalies using passive probes", IEEE INFOCOM, 2007.

[5] R. Kumar , J. Kaur, and "Efficient Beacon Placement for Network Tomography", IMC, 2004.

[6] BRITE, [Online]. Available: http://www.cs.bu.edu/brite/

[7] Cplex, [Online]. Available: http://www.ilog.com/products/cplex.