LEISURE: A Framework for Load-Balanced Network-Wide Traffic Measurement

†Chia-Wei Chang, ‡Guanyao Huang, †Bill Lin, ‡Chen-Nee Chuah
†University of California, San Diego, ‡University of California, Davis

ABSTRACT
Network-wide traffic measurement is of interest to network operators to uncover global network behavior for the management tasks of traffic accounting, debugging or troubleshooting, security, and traffic engineering. Increasingly, sophisticated network measurement tasks such as anomaly detection and security forensic analysis are requiring in-depth fine-grained flow-level measurements. However, performing in-depth per-flow measurements (e.g., detailed payload analysis) is often an expensive process. Given the fast-changing Internet traffic landscape and large traffic volume, a single monitor is not capable of accomplishing the measurement tasks for all applications of interest due to its resource constraint. Moreover, uncovering global network behavior requires network-wide traffic measurements at multiple monitors across the network since traffic measured at any single monitor only provides a partial view and may not be sufficient or accurate. These factors call for coordinated measurements among multiple distributed monitors.

In this paper, we present a centralized optimization framework, *LEISURE* (**L**oad-**E**qual**I**zed mea**SURE**ment), for load-balancing network measurement workloads across distributed monitors. Specifically, we consider various load-balancing problems under different objectives and study their extensions to support different deployment scenarios. We evaluate LEISURE via detailed simulations on Abilene and GEANT network traces to show that LEISURE can achieve much better load-balanced performance (e.g., 4.75X smaller peak workload and 70X smaller variance in workloads) across all coordinated monitors in comparison to naive solution (uniform assignment) to accomplish network-wide traffic measurement tasks.

Categories and Subject Descriptors
G.1.6 [**NUMERICAL ANALYSIS**]: Optimization—*Global optimization, Constrained optimization*; C.2.3 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Operations—*Network management, Network monitoring*

Keywords
Load-balanced, network-wide traffic measurement

1. INTRODUCTION
Accurate traffic measurement is essential to a variety of network management tasks, including traffic engineering (TE), capacity planning, accounting, anomaly detection, and security forensics. Many existing studies focus on improving traffic measurement techniques at a single monitor, including adaptive sampling [1], data streaming [2], and heavy-hitter detection mechanisms [3]. These solutions typically examine packet headers to determine if any statistics need to be collected. While these aggregate traffic volume statistics are sufficient for TE purposes, there is an increasing need for fine-grained flow level measurements to perform accurate traffic classifications for security purposes. For example, deep packet inspection (DPI) allows post-mortem analysis of network events and helps understand the payload properties of transiting Internet traffic. Another solution, Network DVR [4], performs selective flow-based trace collection by matching packets against application-specific signatures.

However, doing fine-grained flow level measurements (e.g., analyzing payload) is often an expensive process that requires dedicated hardware (e.g., TCAMs [5]), specialized algorithms, (e.g., Bloom Filters [6]), or vast storage capacity. Given the fast-changing Internet traffic landscape and large traffic volume, a single monitor is not capable of accomplishing the measurement tasks from all applications of interest due to its resource constraint. This calls for coordinated measurement between multiple distributed monitors. Moreover, network-wide traffic measurement at multiple monitors is also key to uncovering global network behavior since traffic measured at a single monitor only provides partial views and may not be sufficient or accurate. For example, a *global iceberg* [7] may have high aggregate volume across many different monitors, but may not be detectable at any single monitor. Discovering this type of event is important for a number of applications (e.g. detecting DDoS attacks, discovering worms, as well as ensuring SLA compliance).

To perform effective network-wide traffic measurement across multiple distributed monitors, a centralized framework that coordinates measurement responsibilities across different monitors is needed. In today's network, deployed monitors measure traffic completely independently to each other, leading to redundant flow measurements and inefficient use of routers' measurement resources. Sekar et al. [8] proposed CSAMP (Coordinated Sampling), a centralized hash-based packet selection system as a router-level primitive, to allow distributed monitors to measure disjoint sets of traffic without requiring explicit communications, thus eliminating redundant and possibly ambiguous measurements across the network. CSAMP uses an optimization framework to specify the set of flows that each monitor is required to record by considering a hybrid measurement objective that maximizes the total flow-coverage subject to ensuring that the optimal minimum fractional coverage of the task can be achieved. However, both traffic characteristics and measurement tasks can dynamically change over time, coupled with ever-increasing link rates (high traffic volume) and out of consideration to distribute multiple measurement tasks jointly, rendering previously-placed monitors easily overwhelmed if the measurement tasks are not judiciously load-balanced across them, thus leading to entire coordinated measurements failure and wastage of routers' measurement resources. In addition, existing frameworks (e.g., CSAMP) are agnostic to differentiation in the importance of traffic sub-populations or the cost of individual measurement tasks.

We present a new centralized optimization framework called *LEISURE* (**L**oad-**E**qual**I**zed mea**SURE**ment) to address the network measurement load-balancing problem on various realistic scenarios. In contrast to CSAMP whose objective is to maximize the total flow-coverage, LEISURE distributes traffic measurement tasks evenly across coordinated monitors subject to ensuring that the required fractional coverage of those tasks can be achieved. It takes a) routing matrix, b) the topology and monitoring infrastructure deployment and c) measurement requirements of tasks as inputs, and decides which available monitors should participate in each specific measurement task and how much they need to measure to optimize the load-balancing objectives. Ideally the *load-balancing objective* is to have identical workload for all monitors where workload denotes the normalized traffic amount that each monitor measures. In this work, the *load-balancing objective* is mainly defined as two terms: 1) minimizing the variance of workloads across all monitors or 2) minimizing the maximum workload among them. We summarize our contributions as follows:

- We present LEISURE and formulate the optimization problems for network-wide traffic measurements by considering different load-balancing objectives. The optimal solutions are translated into the disjoint sets of required-measured flows that each monitor is assigned to measure. We also propose simple heuristic solutions to compare with and extend LEISURE to incorporate practical scenarios (constraints), i.e., (a) with limited measuring resources at monitors, (b) with limited number of deployed monitors, (c) with multiple routing paths (e.g., ECMP) for each origin-destination (OD)-pair traffic.

- As proof of concept, we perform detailed simulation studies based on Abilene [9] and GEANT [10] network topologies and traces. Our results show that the significant load-balancing improvement (e.g., 4.75X smaller maximum workload and 70X smaller variance in workloads) is achieved by using LEISURE to optimally distribute the measurement tasks across all coordinated monitors when compared with the naive uniform assignments.

- We extend LEISURE and simulation studies to perform optimizations and sensitivity analysis with respect to multiple measurement tasks that exhibit different importance and incur different costs. We show that LEISURE is flexible enough to assign the correct set of measurement tasks for coordinated monitors to optimize measurement utility given limited measuring resources.

The paper is structured as follows. Section 2 outlines related work. Section 3 motivates our load-balancing problem by showing how measurement tasks can be distributed to several coordinated monitors using diversity of intuitions. We present detailed optimization formulations and solution in Section 4, followed by the discussion of extensions in Section 5. Section 6 describes our simulation setup and evaluation results, and Section 7 concludes the paper.

2. RELATED WORK

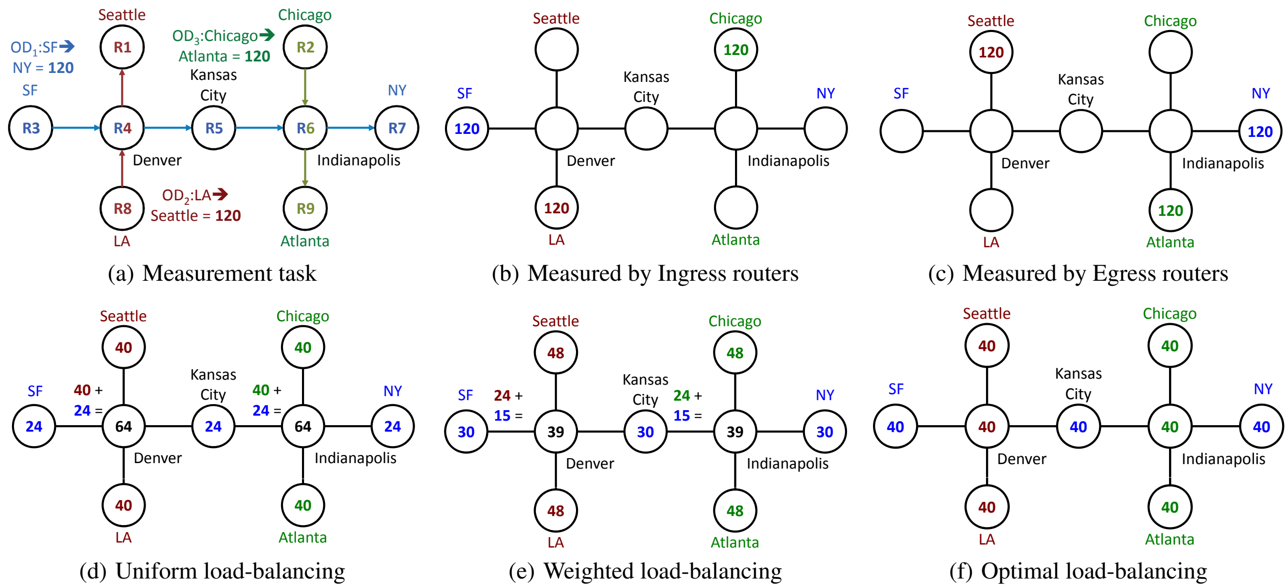
Traffic measurement might involve single point or multiple monitors. Earlier work on traffic measurement has focused on improving single-point measurement techniques, such as sampling approaches [11, 12], estimation of heavy-hitters [3], and methods to channel monitoring resources on traffic sub-populations [13, 14]. Recently, researchers are interested in investigating network-wide traffic measurement problems. In particular, they have demonstrated the benefits of a network-wide approach for traffic engineering [15] and network diagnosis [16].

Network-wide traffic measurement presents more challenges. Previous work on network-wide measurement mostly studied the problem of placing monitors at proper locations to cover all measurement task (routing paths) using as few monitors as possible [17, 18, 19]. Suh et al. [19] first defines utility functions for the sampled traffic and maximizes the overall utility with bounded measurement operation/deployment cost. They propose a two phase approach where they first identify the links that should be monitored and then run a optimization algorithm to set the sampling rates. Cantieni et al. in [20] argues that most ISPs already deploy routers which are equipped with monitoring capabilities (e.g., Netflow [21], Openflow [22]) and these monitoring tools can give greater visibility on the network-wide traffic. Network operators hence can decide whether to turn on these capabilities, and there are potentially hundreds of monitoring points to choose from to achieve network-wide measurements. Based on this assumption, they reformulate the placement problem to decide which monitors should be activated and which sampling rate should be set to achieve a given measurement task with high accuracy and low resource consumption. It performs more rigorous analysis on the convergence of heuristic solutions.

Upon this assumption, our design of LEISURE as a centralized network-wide measurement framework is also encouraged by recent trends in network management. [23, 24] suggest that a centralized network management approach can significantly reduce management complexity and operating costs. [8] showcases that a centralized system that coordinates monitoring responsibilities across multiple routers can significantly increase the flow monitoring capabilities of a network. The global measurement coverage can therefore be improved. In contrast, LEISURE assumes the measurement task can be fulfilled by a given set of numerous monitoring points, and its goal is to optimize the load-balancing objectives by determining which available monitors should participate in each specific measurement task and how much they need to measure instead of solving coverage optimization problem. Also none of the previous work ever considered possible large measurement traffic, multiple measurement tasks with different costs and differentiation in the importance of traffic sub-populations, let alone load balancing among distributed monitors.

3. MOTIVATING EXAMPLE

We first consider the toy example with traffic demands from three OD-pairs: SF$\rightarrow$NY, LA$\rightarrow$Seattle, and Chicago$\rightarrow$Atlanta, each with 120 units of traffic (IP flows) in Fig. 1. Suppose the measurement task imposed by the network operator is to measure all the traffic from these three OD-pairs, one naive approach is to simply always measure the traffic for each OD-pair at the ingress router as shown in Fig. 1(b). The monitors then only need to be placed in SF, LA, and Chicago with measurement traffic as 120 units. Similar to this approach, the traffic for each OD-pair can be measured at the egress router as Fig. 1(c) shown. The monitors instead need to be



(a) Measurement task (b) Measured by Ingress routers (c) Measured by Egress routers

(d) Uniform load-balancing (e) Weighted load-balancing (f) Optimal load-balancing

Figure 1: Different load-balancing approaches for our toy example, which includes three OD-pair traffic as our measurement task (i.e., SF→NY, LA→Seattle, and Chicago→Atlanta, each with 120 units of traffic).

placed in NY, Seattle, and Atlanta with the same measurement traffic. Both of these approaches only need 3 monitors to accomplish the assigned measurement task but with 120 unit measurement traffic.

On the other hand, assume all of these routers are equipped with monitors that are capable of performing the measurement task, our goal is to reduce their maximum measurement traffic by determining a *fraction* of the required measurement traffic to each of these monitors. One simple strategy is to *uniformly* distribute the required measurement traffic of each OD-pair to the monitors along its routing path as depicted in Fig. 1(d). For example, the 120 units of traffic for SF→NY is measured uniformly across monitors placed in SF, Denver, Kansas City, Indianapolis and NY. Each of them takes the measurement responsibility as 24 units. Similarly, the monitors in LA, Denver, Seattle and Chicago, Indianapolis, Atlanta take the measurement responsibility as 40 units for LA→Seattle, and Chicago→Atlanta traffic respectively. The maximum measurement traffic therefore is most likely be the router with the largest number of OD-pairs passing through it (e.g., 64 units of measurement traffic in Denver/Indianapolis).

The other intuitive method distributes the required measurement traffic of each OD-pair to the monitors inverse-proportion-to the traffic passing through them as shown in Fig. 1(e). For example, the traffic passing through SF, Denver, Kansas City, Indianapolis and NY is 120, 240, 120, 240 and 120 respectively. Based on its calculation, SF, Kansas City and NY should measure 30 units of traffic for SF→NY ($\frac{120^{-1}}{120^{-1}+240^{-1}+120^{-1}+240^{-1}+120^{-1}} \times 120$) while Denver and Indianapolis is 15 units. Similarly, the monitors in LA, Seattle, Chicago, Atlanta and Denver, Indianapolis should take the measurement responsibility as 48, 24 units for the traffic LA→Seattle, and Chicago→Atlanta respectively.

Although these two methods achieve significant reduction in the maximum measurement traffic compared to the naive approaches (e.g., 120→64,120→48), it actually can be further reduced to 40 units as shown in Fig. 1(f) by using LEISURE to solve the global load-balancing optimization problem. In this optimal solution, the SF→NY traffic is measured uniformly by only *three* monitors (SF, Kansas City, and NY) instead of five, each with 40 units of traffic while Denver and Indianapolis are *not* involved in the measurement of the SF→NY traffic. This in turn allows the equal splitting of the LA→Seattle traffic and the Chicago→Atlanta traffic across all three routers in each of its respective path, which results in all monitors having the same perfectly load-balanced measurement traffic as 40 units.

It is important to see that the routing path for each OD-pair traffic must overlap, such that the shared monitors can be best utilized by LEISURE to optimally minimize their maximum measurement traffic. If the monitors for measuring each OD-pair traffic are disjoint, there is no opportunity for LEISURE to globally coordinate the overall measurement task since it can only balance the monitors for each OD-pair traffic separately. Therefore the performance of LEISURE in this case degrades as the simple uniform assignments. Next, we are in general interested in finding globally optimal load-balancing solutions by using LEISURE under different network conditions (e.g., topology, traffic demand, routing matrix, etc), measurement objectives (e.g., minimize maximum workload, maximize measurement utility, etc), and resource constraints (e.g., subset of routers are capable of monitoring, some monitors have lower capacities, etc).

4. LEISURE FRAMEWORK

We now present a load-balanced optimization framework to cover network-wide monitoring objectives while respecting router resource constraints. ISPs typically specify their network-wide measurement task in terms of OD-pairs. To cover these measurement assignments, LEISURE needs both the traffic demand and routing information, which are readily available to network operators in [15]. In general, LEISURE is a centralized architecture to allocate disjoint sets of required-measurement flows in OD-pairs for each router by given global network-wide information: a) network topology, monitoring infrastructure deployment, b) traffic demand, routing matrix and c) measurement requirements and the associated cost for each measurement task.

The disjoint sets of required-measurement flows for each router in LEISURE could be implemented by using *hash-based packet selection* in [8] as CSAMP used, a router-level primitive suggested in Trajectory Sampling [25]. Trajectory Sampling assigns all routers in the network a common hash range and each router in the network records the passage for all packets that fall in this common hash range for applications such as fault diagnosis. In contrast, we use hash-based packet selection to assign disjoint hash ranges across multiple routers to ensure the non-overlapping measurement of traffic among monitors as CSAMP. The implementation cost of hash-based packet selection in routers could be found in [8]. Note that both LEISURE and CSAMP use the same hash-based coordination between monitors to implement disjointed flow-measurement. However, our disjoint sets of required-measurement flows for each router are the optimal result which distributes traffic measurement tasks evenly across coordinated routers while in CSAMP, their disjoint flow sets are derived from the output of an optimization framework which aims to maximize the flow-coverage objectives.

The problem formulation builds up from the simplest case in which we assume: 1) the traffic matrix and routing information for the network are given exactly and they change infrequently; 2) flows of each OD-pair follow a single router-level path by OSPF; and 3) there is only one measurement task for every monitor. These constraints are gradually relaxed in Section 5.

4.1 Basic Model

Let *G(V,E)* represent our network topology, where *V* is the set of routers (monitors) and *E* is the set of directed links. Each router V_i $(i = 1 \dots M)$ has two factors to limit its measurement ability: memory and bandwidth. We abstract them into a single resource constraint C_{v_i} $(i = 1 \dots M)$, the number of flows router V_i can measure in a given measurement interval.

An OD-pair, OD_x, represents a set of flows between the same pair of ingress/egress routers for which an aggregated routing placement is given. The set of all $|V| \times |V-1|$ OD-pairs is given by Θ: $OD_x, x \in \Theta$. Φ_x characterizes the traffic demand (IP flows) of the OD-pair $OD_x, x \in \Theta$ in a given measurement interval (e.g., 5 minutes). P_x represents the given routing strategy (router-level path) for every OD-pair $OD_x, x \in \Theta$.

a_x denotes the desired coverage fraction of IP flows of OD_x that is required to measure, which is imposed by the network operator. Therefore the total required measurement traffic (number of flows), β, introduced to all routers is simply a summation of traffic demand per OD-pair times a_x as $\beta = \sum_{x\in\Theta} \Phi_x \times a_x$.

Let d_i^x denote the fraction of traffic demand (IP flows) of OD_x that router V_i samples/measures (i.e., $d_i^x = \frac{\text{measured flows in } \Phi_x}{\Phi_x}$) while L_i denotes the total traffic (number of IP flows) that router V_i measures for all OD-pairs, $OD_x, x \in \Theta$ normalized by β. The summation of L_i for all routers V_i $(i = 1 \dots M)$ then equals 1. We have:

$$\beta = \sum_{x\in\Theta} \Phi_x \times a_x \tag{1}$$

$$L_i = \frac{1}{\beta} \sum_{x:V_i\in P_x} d_i^x \times \Phi_x \qquad \forall i \tag{2}$$

$$\sum_{i=1}^{M} L_i = 1 \tag{3}$$

Our decision variable is d_i^x. The first constraint of d_i^x is that the value of d_i^x is bounded between 0 and 1 as Eq. (4). The second constraint is that the summation of d_i^x along the path P_i for each OD-pair $OD_x, x \in \Theta$ is a_x, as Eq. (5). If router V_i is not in the routing path P_x of OD-pair $OD_x, x \in \Theta$ $(V_i \notin P_x)$, d_i^x is inherently 0. The third constraint is that the measured fraction of β for each monitor V_i should not exceed its measurement ability (resource constraint) C_{v_i} as Eq. (6). Notations are also summarized in Table I.

$$0 \le d_i^x \le 1 \qquad \forall x, i \tag{4}$$

$$\sum_{i:V_i\in P_x} d_i^x = a_x \qquad \forall x \in \Theta \tag{5}$$

$$\sum_{x:V_i\in P_x} d_i^x \times \Phi_x \le C_{v_i} \qquad \forall i \tag{6}$$

4.2 Problem Formulation

We define our load-balancing objective in abstract form α, which can be any term as long as it captures load-balancing performance (i.e., identical workload for all monitors). The overall optimization objective of LEISURE is to minimize α that each router operates within its resource constraint by given parameter a_x, the required fractional coverage per OD-pair imposed by the network operator. In this section, we formulate and study three different optimization problems that correspond to three different load-balancing objective α: *min-VAR*, *min-MAX* and *min-VAR-given-MAX*.

4.2.1 Minimize Variance Problem (min-VAR)

In this problem, we denote α as the variance of L_i across all participating routers[1]. The intuition is that with more even workload L_i for all routers, the variance is smaller (e.g., variance=0 stands for ideal load-balancing objective where $L_i = \frac{1}{M}$ for all M routers). We have:

$$\alpha = VAR(L_i) = \frac{\sum_{i=1}^{M}(L_i - \bar{L})^2}{M} \tag{7}$$

$$\bar{L} = \frac{1}{M}\sum_{i=1}^{M} L_i = \frac{1}{M} \cdot 1 \tag{8}$$

[1]We use "population variance" instead of "sample variance" as our objective function since we already know the number of monitors m.

Table 1: Notations

Notation	Description
OD_x	represent a set of flows between the same pair of ingress/egress routers
Θ	the set of all $\|V\| \times \|V-1\|$ OD-pairs: $OD_x, x \in \Theta$
Φ_x	characterizes the traffic demand (IP flows) of OD-pair $OD_x, x \in \Theta$
P_x	represents the given routing strategy for OD-pair $OD_x, x \in \Theta$
a_x	the fraction of Φ_x (IP flows) of OD_x that is required to measure
d_i^x	the fraction of Φ_x (IP flows) of OD_x that router V_i measures
β	the total required measurement traffic (number of IP flows)
L_i	the total traffic (number of IP flows) that V_i measured normalized by β
α	load-balancing objective

Table 2: d_i^x for each approach with the toy example shown in Fig. 1

	d_3^1	d_4^1	d_5^1	d_6^1	d_7^1	d_1^2	d_4^2	d_8^2	d_2^3	d_6^3	d_9^3	$MAX(L_i)$	$VAR(L_i)$	# of monitors	Decision
LB(ingress)	1	0	0	0	0	0	0	1	1	0	0	120/360	0.025	3	local
LB(egress)	0	0	0	0	1	1	0	0	0	0	1	120/360	0.025	3	local
LB(uniform)	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1/3	64/360	0.00167	9	local
LB(weighted)	1/4	1/8	1/4	1/8	1/4	2/5	1/5	2/5	2/5	1/5	2/5	48/360	0.000484	9	global
LB(optimal)	1/3	0	1/3	0	1/3	1/3	1/3	1/3	1/3	1/3	1/3	40/360	0	9	global

This optimization problem is formulated as:

$$\text{minimize } \alpha = VAR(L_i)$$
$$\text{subject to}$$

$$\frac{1}{\beta} \sum_{x:V_i \in P_x} d_i^x \times \Phi_i = L_i \qquad \forall i \qquad (9)$$

$$\sum_{i:V_i \in P_x} d_i^x = a_x \qquad \forall x \in \Theta \qquad (10)$$

$$\sum_{x:V_i \in P_x} d_i^x \times \Phi_x \le C_{v_i} \qquad \forall i \qquad (11)$$

$$0 \le d_i^x \le 1 \qquad \forall x, i \qquad (12)$$

4.2.2 Minimize Maximum Problem (min-MAX)

In this problem, we denote α as the maximum value of L_i across all routers:

$$\alpha = MAX(L_i) \qquad i = 1 \dots M \qquad (13)$$

The intuition is that when LEISURE keeps minimizing the maximum value of L_i for all monitors by adjusting decision variables d_i^x, other smaller L_i will increase, eventually they will reach some equilibrium state that no more adjustments it can do to lower the $MAX(L_i)$ without increasing other L_i above $MAX(L_i)$. The problem formulation shares the same constraints as min-VAR problem, Eq.(9-12), except that the objective function is different: minimize $\alpha = MAX(L_i), i = 1 \dots M$.

4.2.3 Minimize Variance with Max-Constraint Problem (min-VAR-given-MAX)

This problem involves two phases. In the first step, we formulate the *min-MAX* problem given in Section 4.2.2 to find the minimum achievable maximum value L_{max} (L_{max}= minimized $MAX(L_i), i = 1 \dots M$) for all routers to cover the total required-measurement IP flows, β. Then we seek for any opportunity to further re-distribute the measurement task (workload) evenly within this constraint. Therefore in the second step, we introduce additional constraints to the *min-VAR* problem given in Section 4.2.1 to limit the L_i for each router V_i to be at most L_{max}. We then minimize the variance of L_i across all routers. Specifically, we only need to introduce the following constraint to the *min-VAR* problem:

$$L_i = \frac{1}{\beta} \sum_{x:V_i \in P_x} d_i^x \times \Phi_x \le L_{max} \qquad \forall i \qquad (14)$$

Therefore the *min-VAR-given-MAX* problem actually combines the *min-VAR* and *min-MAX* problems.

4.3 Optimal/Heuristic Solutions

We seek for the optimal d_i^x assignments for the above three problems. There is a variety of optimization tools that we can leverage. Specifically, the optimal solutions can be found by using a Quadratic Programming (QP) formulation for the *min-VAR* problem and a Linear Programming (LP) formulation for the *min-MAX* problem. The combined problem, *min-VAR-given-MAX*, can be solved in a two-phase manner by using LP first and QP follows. We refer these three optimal solutions of LEISURE as LB(min-VAR), LB(min-MAX), and LB(min-VAR-given-MAX), respectively.

Besides the optimal solutions, we introduce one simple heuristic method called LB(weighted) under the assumption that routers can always fulfill assigned measured tasks (e.g., no resource constraints for all routers in Eq. (6)). LB(weighted) calculates d_i^x in inverse-proportion to the total required-measurement traffic amount (IP flows) passing through router V_i. The rationale behind it is that routers with larger required-measurement IP flows passing through should be assigned with less IP flows to measure in order to achieve load-balancing objective. Let β_i denote the total required measurement traffic passing through router V_i, which can be calculated using Eq. (15). The d_i^x assignment for LB(weighted) is formulated as:

$$\beta_i = \sum_{x:V_i \in P_x} \Phi_x \cdot a_x \qquad \forall i \in V \tag{15}$$

$$d_i^x = \frac{\frac{1}{\beta_i}}{\sum_{i:V_i \in P_x} \frac{1}{\beta_i}} \times a_x \qquad \forall x, i \tag{16}$$

Although LB(weighted) does not necessarily lead to the optimal solution, its computation time is very fast comapred to the time required to solve QP or LP optimization problems for LB(min-VAR), LB(min-MAX), and LB(min-VAR-given-MAX). In Section 6, we compare their load-balacning performances also with the following three simple naive strategies:

1. LB(ingress): the required measurement traffic $\Phi_x \cdot a_x$ for each OD-pair $OD_x, x \in \Theta$ is only measured at ingress routers.
2. LB(egress): the required measurement traffic $\Phi_x \cdot a_x$ for each OD-pair $OD_x, x \in \Theta$ is only measured only at egress routers.
3. LB(uniform): the required measurement traffic $\Phi_x \cdot a_x$ for each OD-pair $OD_x, x \in \Theta$ is measured evenly across the routers on its routing path P_x.

Table 2 summarizes the corresponding d_i^x for each approach with the toy example presented in Fig. 1. In this example, LB(min-VAR), LB(min-MAX), and LB(min-VAR-given-MAX) all have the same optimal load-balancing performance (i.e., $MAX(L_i) = \frac{40}{360}$ and $VAR(L_i) = 0$), which we denote as LB(optimal). In comparison, LB(ingress) and LB(egress) have poorest load-balancing performance but with least number of deployed monitors. LB(uniform) outperforms them but needs more monitors (e.g., 9 instead of 3 monitors in our toy example). LB(weighted) and LB(optimal) which consider global required measurement traffic can have better load-balancing performance compared to the local approaches (e.g., LB(ingress), LB(egress) and LB(uniform)), where LB(optimal) has the optimal load-balancing performance but needs much more computation time.

5. EXTENSIONS

In this section, we extend previous formulations to cover some practical scenarios, including 1) only a subset of routers are deployed with monitors; 2) traffic from each OD-pair follows multiple paths (e.g., ECMP: equal cost multiple path); and 3) multiple measurement tasks with different measurement costs and importance factors.

5.1 Measurement with Limited Monitors

In practice, not every router is capable of measurement. Suppose K out of the M routers are deployed with monitors and have measurement capability. We assume each OD-pair $OD_x, x \in \Theta$ has at least one router on its routing path P_x which is capable of measurement to fulfill the measurement tasks imposed by the network operator.

Our formulation includes two scenarios. In the first case, we assume that the K monitors have been deployed and fixed. Our goal is to distribute required measurement tasks to these limited K routers. It can be simply solved by changing the routing index P_x. We exclude router V_i from P_x if it is unable to measure as $P_x^* = P_x - \{V_i\}$ for all OD-pair $OD_x, x \in \Theta$. Variance calculation should also be modified since we now have K monitors instead of M. Other formulations remain the same except that P_x are replaced by P_x^* in all constraints.

$$P_x^* = P_x - \{V_i\}, \text{if } V_i \text{ is unable to measure} \tag{17}$$

$$VAR(L_i) = \frac{\sum_{i=1}^{K} (L_i - \bar{L})^2}{K} \tag{18}$$

$$\bar{L} = \frac{1}{K} \sum_{i=1}^{K} L_i = \frac{1}{K} \tag{19}$$

In the second case, the location of K monitors have not been decided and they are flexible to be deployed in any router. This problem includes not only the *distribution* of measurement tasks, but also the *placement* of monitors. To solve this problem, we introduce additional decision variables u_i, where $u_i = 1$ if router V_i is selected to deploy a monitor, and $u_i = 0$ otherwise. The summation of u_i is therefore K. We assume every monitor has identical limited measurement ability (resource constraint) as C_m. The problem is formulated below with load-balancing objective as either $\alpha = MAX(L_i)$ or $\alpha = VAR(L_i)$ by substituting Eq. (18). Note that it is no longer an LP/QP problem since $u_i, i \in V$ are Boolean variables.

$$\text{minimize } \alpha$$

$$\text{subject to}$$

$$\frac{1}{\beta} \sum_{x:V_i \in P_x} d_i^x \times \Phi_x \times u_i = L_i \qquad \forall i \tag{20}$$

$$\sum_{i:V_i \in P_x} d_i^x \times u_i = a_x \qquad \forall x \in \Theta \tag{21}$$

$$\sum_{x:V_i \in P_x} d_i^x \times \Phi_x \times u_i \le C_m \qquad \forall i \tag{22}$$

$$\sum_{i=1}^{M} u_i = K \tag{23}$$

$$0 \le d_i^x \le 1 \qquad \forall x, i \tag{24}$$

$$u_i \in \{0, 1\} \qquad \forall i \tag{25}$$

5.2 Multi-Path Routing

All the sections above have assumed single-path routing (e.g., OSPF). In this section, we extend our work to support "load-balancing" of measurement tasks in the case of multi-path routing (e.g., ECMP). Since ECMP enables routers to make forwarding decisions on a per IP-flow basis rather than on a per-packet basis, packets for a single flow will still follow one path.

Our formulation treats each of the different paths as a distinct virtual OD-pair with different portions of the origin traffic demand. Suppose each OD-pair $OD_x, x \in \Theta$ has N_x routing paths, denoted as P_{xh} ($h = 1 \ldots N_x$) with total traffic demand Φ_x. We create virtual OD-pairs OD_{xh} for each path P_{xh} ($h = 1 \ldots N_x$) of OD-pair $OD_x, x \in \Theta$ with traffic demand Φ_{xh} where $\sum_{h=1}^{N_x} \Phi_{xh} = \Phi_x, \forall x$. We also let a_{xh} denote the given fraction of Φ_{xh} that is required to be measured for each virtual OD-pair OD_{xh} where $\sum_{h=1}^{N_x} a_{xh} = a_x, \forall x$. d_i^{xh} denotes the fraction of Φ_{xh} that router V_i measures for each virtual OD-pair OD_{xh}. The problem can be formulated below. In this formulation, d_i^{xh} are the decision variables. L_i and a_{xh} can in turn be calculated as functions of d_i^{xh}. α

can still be defined according to different optimization criteria.

$$\text{minimize } \alpha$$
$$\text{subject to}$$
$$\sum_{x\in\Theta}\sum_{h=1}^{N_x}\Phi_{xh}\times a_{xh}=\beta \tag{26}$$
$$\frac{1}{\beta}\sum_{x:V_i\in P_{xh}}\sum_{h=1}^{N_x}d_i^{xh}\times\Phi_{xh}=L_i \qquad \forall i \tag{27}$$
$$\sum_{i:V_i\in P_{xh}}d_i^{xh}=a_{xh} \qquad \forall h,x \tag{28}$$
$$\sum_{x:V_i\in P_{xh}}\sum_{h=1}^{N_x}d_i^{xh}\times\Phi_{xh}\le C_{v_i} \qquad \forall i \tag{29}$$
$$0\le d_i^{xh}\le 1 \qquad \forall h,x,i \tag{30}$$

5.3 Measurement with Multiple Tasks

Until now, we have assumed a single measurement task/function with identical unit cost at every router. In practice, traffic measurement may involve multiple tasks with different measurement cost factors (e.g., DPI is much more resource-intensive than say counting). It is important that we evenly distribute measurement tasks to monitors in this setting. Meanwhile, in some fringe cases, different measurements might compete for limited resources. It is also important to study how they cooperate to achieve better global measurement.

Therefore we have two optimization objectives: 1) minimize the maximum value of L_i for all routers ($i = 1 \dots M$) from load-balancing perspective; 2) maximize the aggregated measurement utility across all measurement tasks. This joint optimization problem involves two phases. In the first step, we use the *min-MAX* problem formulation given in Section 4.2.2 to find the minimum achievable maximum value L_{max} to fulfill every requested measurement task for all routers by temporarily ignoring routers' measurement capabilities (resource constraints). In the second step, we introduce θ_i to reflect the resource constraints for all routers by limiting their L_i to not exceed $\theta_i \times L_{max}$ as $L_i \le \theta_i \times L_{max}, i = 1 \dots M$ where $0 \le \theta_i \le 1$. The more severe the resource constraint is (i.e., with smaller C_{V_i}), the lower the θ_i will be while $\theta_i = 1$ means no resource constraint for router V_i. We then maximize the measurement utility for all tasks under limited resource constraints and load-balancing conditions.

We assume there are in total ζ measurement tasks. Each task, denoted as t ($t = 1 \dots \zeta$), is characterized by its measurement cost C^t. Let a_{xt} denote the given fraction of Φ_x that is required to be measured for each measurement task t ($t = 1 \dots \zeta$) per OD-pair $OD_x, x \in \Theta$. We assume single path routing for every OD-pair $x \in \Theta$ and all routers are capable of processing every measurement task. Our first optimization problem is to evenly distribute the measurement tasks/costs across all routers where the measurement capabilities (resource constraints) of all routers are temporarily ignored. We choose the load-balancing objective as α =MAX(L_i) and the problem is formulated below.

$$\text{minimize } \alpha = \text{MAX}(L_i)$$
$$\text{subject to}$$
$$\sum_{x\in\Theta}\Phi_x\times\sum_{t=1}^{\zeta}a_{xt}\times C^t=\beta \tag{31}$$
$$\frac{1}{\beta}\sum_{x:V_i\in P_x}\Phi_x\times\sum_{t=1}^{\zeta}d_i^{xt}\times C^t=L_i \qquad \forall i \tag{32}$$
$$\sum_{i:V_i\in P_x}d_i^{xt}=a_{xt} \qquad \forall x,t \tag{33}$$
$$0\le d_i^{xt}\le 1 \qquad \forall x,t,i \tag{34}$$

After the optimal minimum achievable maximum workload L_{max} is found for every router (with no resource constraint) to cover all measurement tasks (L_{max}= minimized $MAX(L_i), i = 1 \dots M$), we next consider that routers have their own resource constraint C_{V_i} which may make their $L_i < L_{max}$ and fail partial measurement tasks. Let θ_i denote the fraction of L_{max} to reflect the resource constraint C_{V_i} for router V_i ($i = 1 \dots M$) as $\theta_i = \min(\frac{C_{V_i}}{L_{max}\cdot\beta}, 1)$ where $0 \le \theta_i \le 1$ [2] We introduce a new constraint for all routers to bound their L_i by $\theta_i \times L_{max}$ as[3]:

$$L_i\le\theta_i\times L_{max} \qquad \forall i \tag{35}$$

Under this constraint, we study how different measurement tasks are assigned with proper portion of resources such that the overall measurement utility is maximized. Let I^t denote the importance factor for each measurement task t ($t = 1 \dots \zeta$) and G denote the ideal aggregated measurement utility weighted by I^t for all measurement tasks without considering resource constraints at routers, $G = \sum_{x\in\Theta}\Phi_x \times \sum_{t=1}^{\zeta}$. G_i denotes the total measurement utility that router V_i gets for all measurement tasks normalized by G, $G_i = \frac{1}{G}\sum_{x:V_i\in P_x}\Phi_x \times \sum_{t=1}^{\zeta} d_i^{xt} \times I^t$. The optimization problem can be formulated as follows, with d_i^{xt} as the decision variables:

$$\text{maximize } \sum_{i=1}^{M}G_i$$
$$\text{subject to}$$
$$\sum_{x\in\Theta}\Phi_x\times\sum_{t=1}^{\zeta}a_{xt}\times C^t=\beta \tag{36}$$
$$\frac{1}{\beta}\sum_{x:V_i\in P_x}\Phi_x\times\sum_{t=1}^{\zeta}d_i^{xt}\times C^t=L_i \qquad \forall i \tag{37}$$
$$\theta_i\times L_{max}\ge L_i \qquad \forall i \tag{38}$$
$$\sum_{i:V_i\in P_x}d_i^{xt}\le a_{xt} \qquad \forall x,t \tag{39}$$
$$0\le d_i^{xt}\le 1 \qquad \forall x,t,i \tag{40}$$

[2] θ_i=1 implied that there is no resource constraint on router V_i since the traffic amount it measured is less than its resource constraint: $L_i \times \beta \le L_{max} \times \beta \le C_{V_i}$.

[3] By substituting θ_i=min($\frac{C_{V_i}}{L_{max}\cdot\beta}, 1$) with Eq. 32 into Eq. 35, the traffic amount, $\sum_{x:V_i\in P_x}\Phi_x \times \sum_{t=1}^{\zeta} d_i^{xt} \times C^t$, that router V_i measured is always less than its resource constraint (C_{V_i}).

The value of normalized objective function, $\sum_{i=1}^{M} G_i$, is always in the range as $0 < \sum_{i=1}^{M} G_i \leq 1$. For the case when θ_i=1 $\forall i$, $\sum_{i=1}^{M} G_i = 1$ which means all the required measurement tasks can be satisfied ($\sum_{i:V_i \in P_x} d_i^{xt}$=$a_{xt}$ $\forall x, t$) and the aggregated measurement utility is maximum since there is no resource constraints on all routers. However, with the resource constraints (θ_i decreases), only a subset of measurement tasks can be fulfilled, and the goal of the above formulation is to maximize the global measurement utility and maintain the load-balancing conditions simultaneously.

6. PERFORMANCE EVALUATION

We evaluated the performance of LEISURE with three optimal solutions for different load-balancing objectives (i.e., LB(min-VAR), LB(min-MAX) and LB(min-VAR-given-MAX)) in various realistic scenarios on two separate real, large point-of-presence(PoP)-level backbone networks: Abilene [9] and GEANT [10]. We also compare them with several simple heuristic approaches, namely LB(ingress), LB(egress), LB(uniform), and LB(weighted). Our starting point is to conduct a preliminary evaluation on the basic model in Section 6.2 based on three assumptions: (1) all routers are equipped with monitors that are capable of performing the measurement task, (2) traffic from each OD-pair has a single router-level path by OSPF and (3) there is only one measurement task. We relax these assumptions in Section 6.3 and Section 6.4 to show our load-balancing ability. Section 6.5 presents our load-balancing and measurement utility maximizing results for the scenario of multiple measurement tasks with different cost and importance factors.

6.1 Experimental Setup

We use two real datasets from the Abilene [9] and GEANT networks [10], both of which have been studied and discussed in the research literature. Their data sets are publicly available, including network topology, routing information. Based on these available data sets, we implemented a flow-based trace-driven simulation to conduct our evaluations. For both networks, we use the real traffic matrices provided by a third party [26]. The traffic matrix data sets for the Abilene network are available at [27], and the traffic matrix data sets of the GEANT network are available at [28].

Abilene: A public academic network in the U.S. with 11 nodes interconnected by OC192, 10 Gbits/s links. The traces we use were collected from April 22-26, 2004. The routers in ATLA, CHIN, DENV, HSTN, IPLS, KSCY, LOSA, NYCM, SNVA, STTL and WASH are denoted as $R_0, R_1, \cdots, R_{10}$ respectively.

GEANT: It connects a variety of European research and education networks. Our experiments were based on the December 2004 snapshot available at [29], which consists of 23 nodes and 74 links varied from 155 Mbits/s to 10 Gbits/s. The traces we use were collected from April 11-15, 2004.

The traffic matrix we use consists of demands for every OD-pair within a certain time interval (5 mins for Abilene and 15 mins for GEANT). We construct OD-pairs by considering all possible pairs of PoPs and calculate their shortest-path routes. In brief, these traffic matrices are derived from flow information collected at key locations of the network, and is transformed into the demand rate for each OD-pair based on the control plane information.

In the following sections, we assume our target is to measure all traffic (i.e., $a_x = 1, \forall x \in \Theta$). Therefore the workload L_i for router R_i ($i = 1 \ldots M$) is defined as the traffic amount that router R_i measured normalized by the total traffic demand. Theoretically, the ideal load-balancing workload L_i for M monitors is $\frac{1}{M}$. However, it might be unachievable due to routing limitations from TE or resource constraints on monitors.

6.2 Basic Load-Balancing Comparison

In this section, we compare the load-balancing performance of all approaches on two assumptions (ubiquitous monitors and single path routing). The load-balancing performance are compared mainly with respect to two metrics: (1) the maximum value of each monitor's measurement workload in the entire network, namely MAX(L_i), and (2) the variance of workloads across all monitors, namely VAR(L_i).

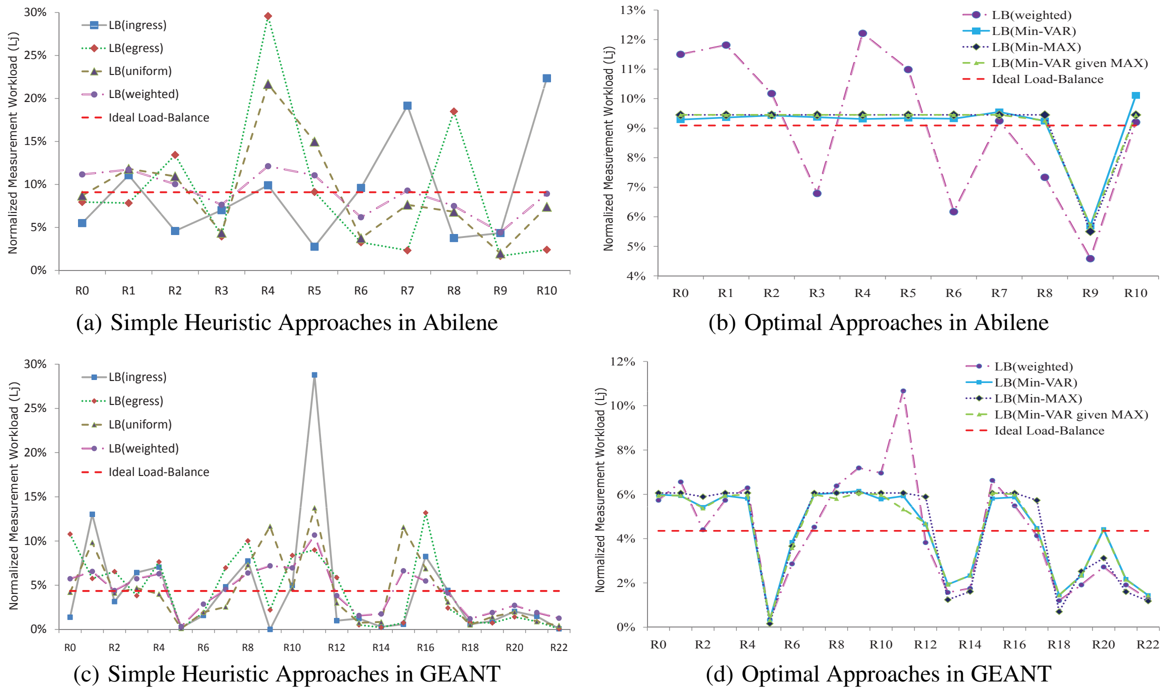
Table 3 compares MAX(L_i) of all monitors for different approaches. For GEANT, our optimal load-balancing solutions can reduce MAX(L_i) by a factor of 4.75X(=$\frac{28.79\%}{6.06\%}$) when compared to the naive approach of LB(ingress) and 2.27X(=$\frac{13.73\%}{6.06\%}$) when compared to LB(uniform). Similar gains can be seen in the results for Abilene as well. Fig. 2 plots in more details the L_i values of 11 monitors and 23 monitors for different load-balancing approaches in Abilene and GEANT networks respectively.

Another relative performance measure is to see how close the maximum workloads are in comparison to the ideal load-balancing case of $\bar{L} = \frac{1}{M}$, as given by Eq.(8). For Abilene and GEANT, the ideal $\bar{L}$ is 9.09%(=$\frac{1}{11}$) and 4.35%(=$\frac{1}{23}$), respectively. However, the MAX(L_i) of LB(ingress) for Abilene and GEANT are 19.16% and 28.79%, respectively, which are 2.11X and 6.62X worse than the ideal case. For simple heuristic approaches, they still have large MAX(L_i) values compared to the ideal case: e.g., 21.67% (2.3X worse) for LB(uniform) in Abilene and 10.67% (2.4X worse) for LB(weighted) in GEANT. On the other hand, our three optimal load-balancing solutions presented in Fig. 3 and Table 3 perform very close to the theoretical ideal case: 9.45%, 10.11%, and 9.45% for LB(min-VAR), LB(min-MAX), and LB(min-VAR given MAX), respectively, as compared to the ideal case of 9.09% for Abilene. Similarly, our three optimal solutions are 6.06%, 6.15%, and 6.06%, respectively, as compared to the ideal case of 4.35% for GEANT.

Table 4 compares VAR(L_i) across all monitors for different approaches. For Abilene, our optimal load-balancing solutions can reduce VAR(L_i) by a factor of 70X(=$\frac{0.007366}{0.000105}$) when compared to the naive approach of LB(egress), and over 30X(=$\frac{0.003158}{0.000105}$) when compared to LB(uniform). Similar improvements in variance can be seen for GEANT as well.

To better understand why our optimal solutions can achieve more evenly distributed measurement load, we use traffic from only five OD-pairs in Abilene[4] to show the detailed load assignment in Fig. 3 (WAS-DNV, NYC-HST, DNV-IPL, CHI-LOS and ATL-STT with 66.5 MB, 44.9 MB, 44.6 MB, 19.8 MB and 11.7 MB, respectively). In Fig. 3(a), although LB(uniform) distributes each OD-pair traffic to all monitors in the path uniformly (e.g., WAS-DNV with 6 monitors), the aggregated workload for overall measurement task in each monitor is still unbalanced (e.g., L_i for all routers R_i ($i = 1 \ldots 10$) are distributed between 1% to 17%). LB(weighted) in Fig. 3(b) improves the load-balancing performance due to the global view it has but still load-balanced poorly (e.g., L_i distributed between 4% to

[4]The notations of these OD-pairs and their routing information could be found in [9], [27].



(a) Simple Heuristic Approaches in Abilene (b) Optimal Approaches in Abilene

(c) Simple Heuristic Approaches in GEANT (d) Optimal Approaches in GEANT

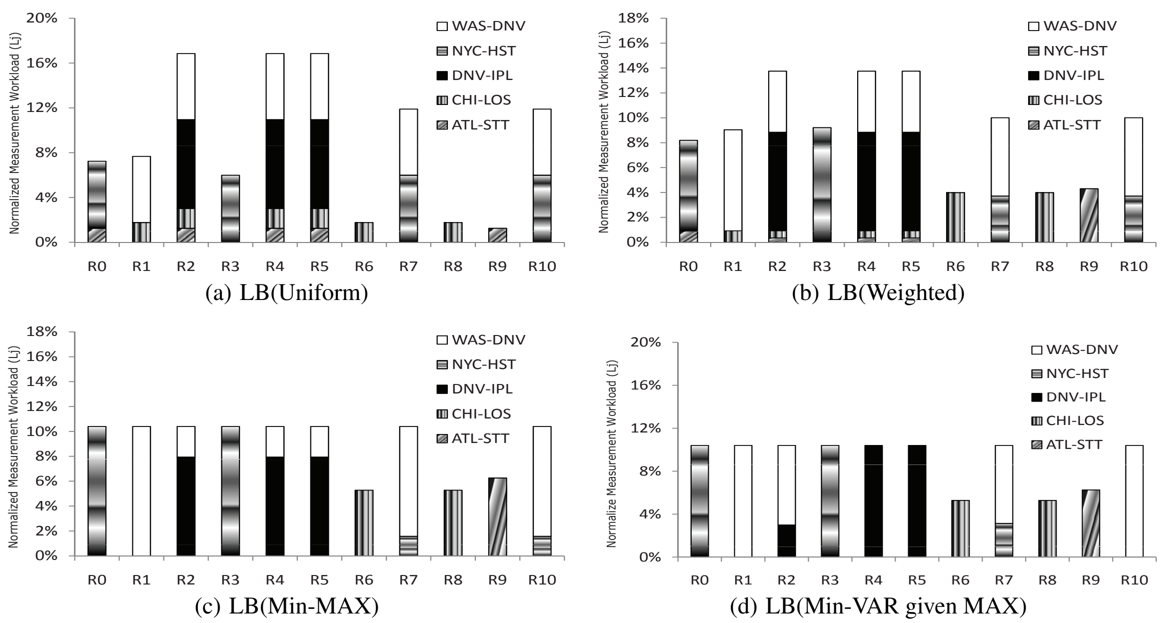
Figure 2: Measurement load distribution for different approaches in Abilene and GEANT

Table 3: Comparisons on Maximum value of L_i

Network	Naive Approaches ingress	Naive Approaches egress	Heuristic Approaches uniform	Heuristic Approaches weighted	Optimal Load-Balancing min-VAR	Optimal Load-Balancing min-MAX	Optimal Load-Balancing min-VAR given MAX
Abilene	19.16%	29.59%	21.67%	12.12%	9.45%	10.11%	9.45%
GEANT	28.79%	13.19%	13.73%	10.67%	6.06%	6.15%	6.06%

Table 4: Comparisons on Variance of L_i

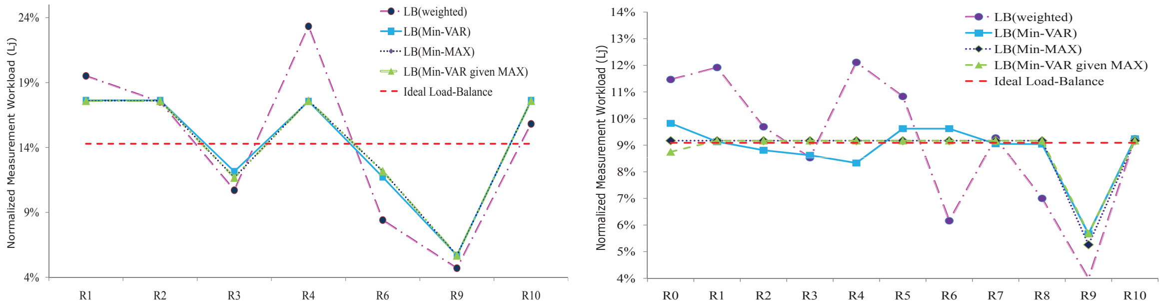
Network	Naive Approaches ingress	Naive Approaches egress	Heuristic Approaches uniform	Heuristic Approaches weighted	Optimal Load-Balancing min-VAR	Optimal Load-Balancing min-MAX	Optimal Load-Balancing min-VAR given MAX
Abilene	0.004107	0.007366	0.003158	0.000602	0.000105	0.000131	0.000105
GEANT	0.003978	0.001626	0.001594	0.000662	0.000378	0.000495	0.000378



(a) LB(Uniform) (b) LB(Weighted)

(c) LB(Min-MAX) (d) LB(Min-VAR given MAX)

Figure 3: Detailed Abilene results for five OD-pairs. Optimal solutions allow nodes to be excluded from measurement if they are already overloaded.



(a) With Limited 7 out of 11 Monitors

(b) With Multiple Paths per OD-pair

Figure 4: Measurement load distribution for different Approaches within different assumptions in Abilene.

14%). In contrast, the optimal solutions can achieve much better load-balancing performance (e.g., L_i distributed between 5.5% to 10.5%) by excluding some monitors from measuring certain OD-pair traffic (e.g., R4 and R5 do not measure traffic for WAS-DNV OD-pair in Fig. 3(d)).

6.3 Limited Measurement Monitors

In this section, we relax our first assumption to the case that only a fraction of routers have measurement capability. We exclude monitors deployed in R_0, R_5, R_7 and R_8 from measurement in Abilene and distribute the measurement task to the remaining 7 monitors[5]. We omit naive approaches and focus on heuristic and optimal approaches in Fig. 4(a). Compared with ubiquitous case in Fig. 2(a), the ideal load-balancing workload is increased from 9.09% to 14.29%. For LB(min-VAR), LB(min-MAX) and LB(min-VAR given MAX), the MAX(L_i) is only increased from 9.67% to 17.61%. However, for heuristic approaches, MAX(L_i) increased from 12.12% to 23.33% for LB(weighted), and 21.67% to 35.86% for LB(uniform). In contrast, we observe that our proposed optimal solutions only increased 7.94% workload for MAX(L_i), which are close to 5.2% for the theoretical ideal case and are much better than 11.21% for LB(weighted) and 14.19% for LB(uniform).

6.4 Multiple Paths per OD-pair

Here we relax our second assumption to allow multi-path routing (e.g,. ECMP) for each OD-pair in Abilene network. In Fig. 4(b), our proposed optimal solutions and heuristic approaches all have better load-balancing performance when applied in multi-path routing case compared to the single path routing. The rationale behind this is that with more overlaps in monitors/paths, LEISURE has more freedom (e.g., d_i^{xh} in Eq. (30)) to optimally load-balance the workloads across the participating monitors. The VAR(L_i) in multi-path case for LB(min-VAR), LB(min-MAX) and LB(min-VAR-given-MAX) is 0.0000917, 0.0000982 and 0.0000917, respectively while in the single path case is 0.000105, 0.000131 and 0.000105 in Fig. 2(b).

6.5 Multiple Measurement Tasks

In this section, we examine our two-phase solution to the formulated joint optimization problem described in Section 5.3. Assume we have two measurement tasks with cost factor ratio C^1:C^2 and

[5]The reason to choose those 4 excluded routers is to maintain the fact that at least one capable monitor in each OD-pair's route to fulfill the measurement tasks imposed by the network operator.

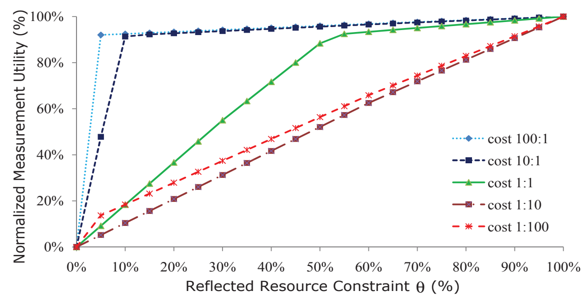


Figure 5: Two measurement tasks with different cost ratio and fixed importance ratio as 1:10.

importance factor ratio I^1:I^2. Let θ reflect the identical resource constraint C_m for all routers V_i ($i = 1 \dots M$) and represented as the fraction of L_{max}, the maximum workload routers can achieve derived from Eq. (31) to (34) where $0 \leq \theta \leq 1$. Fig. 5 presents the result of our normalized measurement utility under different setup of C^1:C^2 (e.g., from 100:1 to 1:100) and fixed I^1:I^2=1:10 by changing resource constraint θ from 100% to 0%. Note that without resource constraint (i.e., $\theta = 100\%$), the normalized measurement utility LEISURE can achieve is always 1.0 (cover all measurement tasks).

As observed, if C^1:C^2 is directly proportional to I^1:I^2, the measurement utility decreases linearly when the resource constraint becomes severe (e.g., lower θ). On the other hand, if C^1:C^2 is inversely proportional to I^1:I^2, the measurement utility will not drop significantly until θ is extremely low (e.g., C^1:C^2=100:1 with I^1:I^2=1:10). This is because the optimal solution for Eq. (36) to (39) will let monitors always first fulfill the measurement request from the task with lower cost and higher importance. The other observation is that when C^1:C^2=1:1 and I^1:I^2=1:10, LEISURE can still remain 90% of the measurement utility as in ideal case (i.e., without resource constraint) by using only half of the routers' resources (e.g., θ drops to 50%(=1/(1+1))). These results suggest that our framework can intelligently distribute measurement tasks for better load-balancing under resource constraints, while the overall measurement utility can still be preserved at a high level.

7. CONCLUSION

In this paper, we proposed an optimization framework for load-balancing network-wide traffic measurements across coordinated monitors in the network. This is an important problem because individual monitors are not capable of accomplishing the measurement tasks for all applications of interest due to its resource constraint, particularly resource-intensive measurement tasks such as those requiring deep packet inspection. Further, to uncover global network behavior, there is an inherent need to coordinate measurements among monitors distributed across the networks since the visibility of each monitor is only limited to the traffic that passes through it. Therefore, these distributed monitors can be coordinated for both coverage and optimized resource utilization. Based on our simulation measurement studies using the Abilene and GEANT networks, we found that our load-balancing optimization framework LEISURE can achieve up to 4.75X smaller maximum measurement workload and 70X smaller variance in workloads across all coordinated monitors. The distributed LEISURE algorithm for load balancing problem is deferred as our future work.

8. REFERENCES

[1] "Sampling and Filtering Techniques for IP Packet Selection," Internet Draft draft-ietf-psamp-sample-tech-11.txt, Work in Progress, July 2008.

[2] A. Kumar, M. Sung, J. Xu, and J. Wang, "Data streaming algorithms for efficient and accurate estimation of flow size distribution," in *Proceedings of ACM SIGMETRICS*, 2004, pp. 177–188.

[3] C. Estan and G. Varghese, "New directions in traffic measurement and accounting: Focusing on the elephants, ignoring the mice," *ACM Transactions on Computer Systems*, vol. 21, no. 3, pp. 270–313, August 2003.

[4] C.-W. Chang, A. Gerber, B. Lin, S. Sen, and O. Spatscheck, "Network dvr: A programmable framework for application-aware trace collection," in *Passive ad Active Measurement Conference*, 2000.

[5] F. Yu, R. H. Katz, and T. V. Lakshman, "Gigabit rate packet pattern-matching using tcam," in *Proceedings of IEEE ICNP*, 2004, pp. 174–183.

[6] S. Dharmapurikar, P. Krishnamurthy, T. Sproull, and J. Lockwood, "Deep packet inspection using parallel bloom filters," in *IEEE Micro*, 2003, pp. 44–51.

[7] G. Huang, A. Lall, C.-N. Chuah, and J. Xu, "Uncovering Global Icebergs in Distributed Monitors," in *IEEE IWQoS*, 2009.

[8] V. Sekar, M. K. Reiter, W. Willinger, H. Zhang, R. R. Kompella, and D. G. Andersen, "CSAMP: A System for Network-Wide Flow Monitoring," in *Proceedings of USENIX NSDI*, 2008.

[9] V. Chan, "Near-term future of the optical network in question?" *IEEE J. Sel. Areas Commun.25*, vol. 25, 2007.

[10] M. Yoo, C. Qiao, and S. Dixit, "Optical burst switching for service differentiation in the next-generation optical Internet," *IEEE Commun. Magazine*, vol. 39, no. 2, pp. 98–104, 2001.

[11] K. C. Claffy, G. C. Polyzos, and H.-W. Braun, "Application of sampling methodologies to network traffic characterization," in *Proceedings of ACM SIGCOMM*, 1993.

[12] N. Hohn and D. Veitch, "Inverting sampled traffic," in *Proceedings of ACM SIGCOMM*, 2003.

[13] A. Ramachandran, S. Seetharaman, N. Feamster, and V. Vazirani, "Fast Monitoring of Traffic Subpopulations," in *Proceedings of ACM Internet Measurement Conference*, 2008.

[14] L. Yuan, C.-N. Chuah, and P. Mohapatra, "ProgME: Towards Programmable Network MEasurement," in *Proceedings of ACM SIGCOMM*, 2007.

[15] Y. Zhang, M. Roughan, N. Duffield, , and A. Greenberg, "Fast Accurate Computation of Large-scale IP Traffic Matrices from Link Loads," in *Proceedings of ACM SIGMETRICS*, 2003.

[16] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing Network-Wide Traffic Anomalies," in *Proceedings of ACM SIGCOMM*, 2004.

[17] M. R. Sharma and J. W. Byers, "Scalable Coordination Techniques for Distributed Network Monitoring," in *Proceedings of Passive and Active network Measurement (PAM) workshop*, April 2005.

[18] C. Chaudet, E. Fleury, I. Lassous, H. Rivano, and M.-E. Voge, "Optimal positioning of active and passive monitoring devices," in *Proceedings of ACM CoNEXT*, 2005, pp. 71–82.

[19] K. Suh and Y. Guo and J. Kurose and D. Towsley, "Locating network monitors: Complexity, heuristics and coverage," in *Proceedings of IEEE INFOCOM*, March 2005.

[20] G. R. Cantieni, G. Iannaccone, C. Barakat, C. Diot, and P. Thiran, "Reformulating the monitor placement problem: Optimal network-wide sampling," in *Proceedings of ACM CoNEXT*, 2006.

[21] "Cisco ios netflow," http://www.cisco.com/warp/public/732/Tech/netflow/.

[22] "The OpenFlow Switch Consortium," http://www.openflowswitch.org.

[23] M. Caesar, D. Caldwell, N. Feamster, J. Rexford, A. Shaikh, and J. V. D. Merwe, "Design and implementation of a Routing Control Platform," in *Proceedings of USENIX NSDI*, 2005.

[24] H. Ballani and P. Francis, "CONMan: A Step Towards Network Manageability," in *Proceedings of ACM SIGCOMM*, 2007.

[25] N. Duffield and M. Grossglauser, "Trajectory Sampling for Direct Traffic Observation," in *Proceedings of ACM SIGCOMM*, 2001.

[26] "Totem, a toolbox for traffic engineering methods, february 2005," http://totem.info.ucl.ac.be.

[27] "Abilene traffic matrices," http://www.cs.utexas.edu/~yzhang/research/AbileneTM.

[28] "Totem geant traffic matrices," http://totem.info.ucl.ac.be/dataset.html.

[29] "Geant topology available at," http://www.geant.net/upload/pdf/GEANT_Topology_12-2004.pdf.