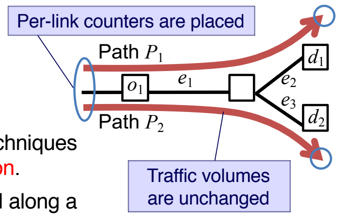
A Proposal of an Efficient Traffic Matrix Estimation under Packet Drops

Kohei Watabe† Toru Mano†† Kimihiro Mizutani†† Osamu Akashi†† Kenji Nakagawa† Takeru Inoue††
† Graduate School of Engineering, Nagaoka University of Technology
†† NTT Network Innovation Laboratories

Background and Objectives

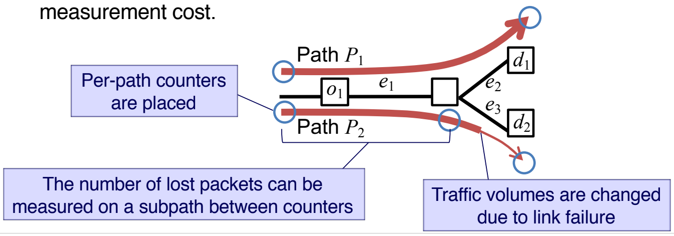
Conventional Traffic Matrix Estimation

- Traffic Matrices (TMs), which specify the traffic volumes between origin-destination pairs in a network, are used by many network engineering tasks.
  - traffic engineering
  - capacity planning
  - anomaly detection etc.
- Conventional TM estimation techniques assume the strict flow conservation.
  - Traffic volumes are unchanged along a path from the origin to the destination.



Objective

- Accurate TMs including volume changes would be very useful for advanced network engineering.
- This paper studies a mathematical model to estimate traffic volumes with their change along a path, by counting the number of packets.
  - The number of counters should be minimized to reduce the measurement cost.



Two baseline approaches

- There are two baseline approaches to solve per-link information using information $\boldsymbol{y}$ of n subpaths.

A : a measurement matrix whose element a_{ij} indicates whether i-th subpath includes j-th link. It is specified by counter placement. $\boldsymbol{x}$

1. Linear algebra approach can exactly determine the traffic volumes with change, but it requires many counters to make TA full-rank.

$$TA\boldsymbol{x} = \boldsymbol{y}$$

$\boldsymbol{x}$: a vector whose j-th element is the packet loss rate on j-th link

$\boldsymbol{y}$: a vector whose i-th element is volume of lost traffic on i-th path

T: an element t_{ii} is a traffic volume of i-th path and the other elements are zero.

2. Boolean algebra approach requires fewer counters, but it only locates failed links without volume change.

$$A\boldsymbol{x} = \boldsymbol{y}$$

$\boldsymbol{x}$: a Boolean vector whose j-th element indicates whether j-th link is failed

$\boldsymbol{y}$: a Boolean vector whose i-th element indicates whether i-th subpath is failed

- This paper establishes a new measurement method that intervenes between the two approaches.
  - The volume change is estimated with error bounds, while it requires counters fewer than the linear algebra approach and close to the Boolean algebra approach.

Formulation

Network Model

- A counter placed in a network is specified by the pair of arc and path.
  - A counter maintains the number of packets transmitted into the arc along the associated path.

$$(e_j, P) \in E \times \mathcal{P}$$

E: the set of arcs in a network; $\mathcal{P}$: the set of feasible paths

- We assume that every arc has either a normal state $l_j \leq \epsilon$ or an abnormal state $\delta \leq l_j$.
  - It is worth noting that our model works without the strong assumption, $\epsilon \ll \delta$, used in [1].
- We assume that every path observes the equal loss rate for the same abnormal arc.
- This paper assumes a single arc failure.

Definition of measurability

- The counter set C is α-measurable if

$$\forall P \in \mathcal{P}, e_j \in P : \alpha \cdot \tau_j \leq \hat{\tau}_j \leq \frac{1}{\alpha} \cdot \tau_j,$$

Lower bound ($\alpha \cdot \tau_j$); Estimator (see below) ($\hat{\tau}_j$); Upper bound ($\frac{1}{\alpha} \cdot \tau_j$)

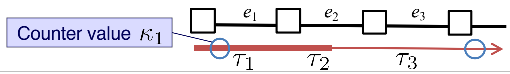
Formulation of the problem

- For given P, $\min_{C \subseteq \mathcal{C}} \{|C| : C \text{ is } \alpha\text{-measurable}\}$

Measurability Theory and Optimization

Estimation

- In our method, an abnormal arc is specified by solving a Boolean equation $A(C)\boldsymbol{x} = \boldsymbol{y}$.
- For a feasible path $P = \{e_1, e_2, \ldots, e_l\}$, our estimator is $\hat{\tau}_j = \kappa_1 \sqrt{(1-\epsilon)^{j-2}}$ if e_j is located on the lower side of the specified abnormal arc, and $\hat{\tau}_j = \kappa_1 \sqrt{(1-\epsilon)^{j-1}}$ otherwise.



Measurability Theorem

- [Theorem] For the above estimator, a counter set C is $(1-\epsilon)^{d-2}$-measurable, if the measurement matrix $A(C)$ is 1-independent and every feasible path P has at least one counter on it.
  - A matrix $A(C)$ is 1-independent if any column vector of $A(C)$ is different from each other and none of them equal to zero vector.

Optimization of Counter Placement

- We initially place counters C_0 at first arcs for every P.
- To satisfy 1-independency, additional counter X that maximizes the following coverage function g is placed repeatedly.
  - The coverage function g is submodular, and the problem is a submodular optimization.

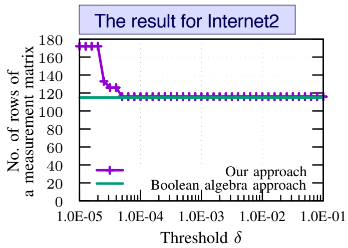
$$g(X) = |\{(j,k) : 0 \leq j < k \leq n, \boldsymbol{a}_j \neq \boldsymbol{a}_k\}|$$

- Subpaths that are longer than $\log_{1-\epsilon}(1-\delta)$ hops are divided by placing counters.
  - Inspecting the counter values, we can tell the longer subpath contains an abnormal arc or not.

Experiments

- Our approach is evaluated using three configuration datasets.
  - Internet2
  - Stanford backbone network
  - Purdue campus network

The result for Internet2



- Though our method can provide the error bounds and traffic volumes, it almost converges to the conventional Boolean technique [1] for $\delta > 5.0 \times 10^{-5}$.

Conclusion

- With the solid theory about the measurability based on Boolean matrices, we developed an optimization algorithm for the minimum counter set.
- Experiments showed the great performance with real datasets.

References

[1] S. Agrawal, K. V. M. Naidu, and R. Rastogi, "Diagnosing Link-level Anomalies Using Passive Probes," in IEEE INFOCOM, 2007, pp. 1757– 1765.