Model-less Approach of Network Traffic for Accurate Packet Loss Simulations

1st Masahiro Terauchi
Graduate School of Engineering
Nagaoka University of Technology
Nagaoka, Niigata, Japan
m_terauchi@kashiwa.nagaokaut.ac.jp

2nd Kohei Watabe
Graduate School of Engineering
Nagaoka University of Technology
Nagaoka, Niigata, Japan
k_watabe@vos.nagaokaut.ac.jp

3rd Kenji Nakagawa
Graduate School of Engineering
Nagaoka University of Technology
Nagaoka, Niigata, Japan
nakagawa@nagaokaut.ac.jp

***Abstract*—It is important to accurately model network traffic when we evaluate Quality of Service (QoS) of networks through simulations. However, for traffic in real networks, it is a tough task to select an appropriate traffic model and tune its parameters. Even if the accurate traffic modeling is achieved, it is also difficult to accurately estimate QoS regarding rare events, such as a packet loss rate in the modern Internet. In this paper, we propose a model-less approach to accurately estimate a packet loss rate through a simulation without directly modeling traffic including real network traffic. We also show the effectiveness of the approach in a simple queueing system as a first step in our development.**

***Index Terms*—Network, Simulation, Importance sampling, Packet loss rate, QoS estimation**

I. Introduction

When we evaluate Quality of Service (QoS) of networks through simulations, accurately modeling traffic of real networks is important and also difficult task. Many traffic models that reflect various characteristics have been proposed by priori works (see [1] and the references therein). However, it is difficult to select an appropriate traffic model and tune its parameters for real network traffic (The double arrows in Fig. 1).

Even if the accurate traffic modeling is achieved, it is also difficult to accurately estimate QoS regarding rare events, such as a packet loss rate in the modern Internet. For accurate estimations of rare events, Importance Sampling (IS) in which a simulation on a condition where the events occur more frequently is performed and the result for the original condition is obtained by the change-of-measure [2]. In the most of researches regarding IS in networks, the change-of-measure is analytically calculated based on the traffic model. They are inapplicable for real network traffic since the applicable traffic models, topology etc. are extremely limited, e.g. Poisson traffic model on a single router etc. (Dotted arrows in Fig. 1).

Though there are a few researches regarding IS in trace-driven simulations without traffic models [3], [4] (Triple arrows in Fig. 1), they cannot be applied for traffic composed of a single flow or correlated flows since they randomize phases of traces. In addition, to the best of our knowledge,

This work was partly supported by JSPS KAKENHI Grant Number JP17K00008 and JP18K18035.

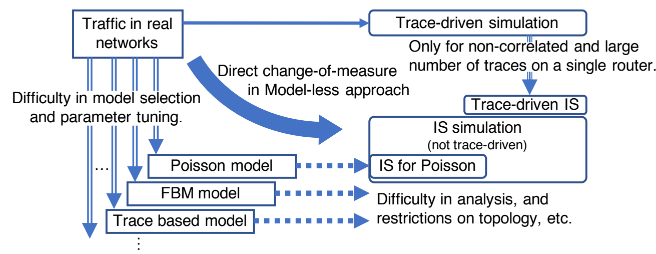

Fig. 1. The model-less approach and the conventional simulations.

no research on trace-driven IS has been published after the researches [3], [4], and there is no prospect of expansion of trace-driven IS to complicated topologies.

In this paper, we propose a model-less approach to accurately estimate a packet loss rate through a simulation without directly modeling traffic, including real network traffic (The bold arrow in Fig. 1). In the model-less approach, a simulation with Poisson traffic model is performed, irrespective of the characteristics of traffic on a network to be evaluated. Naturally, despite the packet loss rate in the network to be evaluated does not correspond to the packet loss rate of the simulation, the estimation result of the loss rate for arbitrary traffic is obtained by a change-of-measure technique based on the model-based IS with a frequency distribution of descritized traffic trace.

II. Notation

In this paper, we express the packet loss rate $\boldsymbol{l} = \{l_i\}_{i\in\mathbb{N}}$ as a probability function $\boldsymbol{l} = f(\boldsymbol{x}(t), \boldsymbol{y}(t))$ of the input and output traffic $\boldsymbol{x}(t) = \{x_i(t)\}_{i\in\mathbb{N}}$ and $\boldsymbol{y}(t) = \{y_i(t)\}_{i\in\mathbb{N}}$, where i denotes flow ID. $x_i(t)$, $y_i(t)$ and $d_i(t)$ are respectively defined as point processes $\sum_{j=1} \delta(t - T_{\mathrm{In}\{i,j\}})$, $\sum_{j=1} \delta(t - T_{\mathrm{Out}\{i,j\}})$ and $\sum_{j=1} \delta(t - T_{\mathrm{Loss}\{i,j\}})$, where $T_{\mathrm{In}\{i,j\}}$, $T_{\mathrm{Out}\{i,j\}}$ and $T_{\mathrm{Loss}\{i,j\}}$ ($\{i,j\} \in \mathbb{N}^2$) are the times when the jth packet in the ith flow inputs, outputs, and losses, respectively. Note that $\delta(t)$ is Dirac delta function. In addition, let $x_{i,T}$ and $y_{i,T}$ be the paths $\{x_i(t)\}_{0\leqq t\leqq T}$ and $\{y_i(t)\}_{0\leqq t\leqq T}$ on $[0, T]$, respectively.

III. Model Based IS

When IS estimates a loss rate on a single router into which a single flow streams, the change-of-measure is performed based on probability density of a path $\omega_{T_{\mathrm{In}\{1,j\}}}$ of the queue length process on $[0, T_{\mathrm{In}\{1,j\}}]$ [2]. Assuming that probability density of an event A in the original condition and the modified condition is $p(A)$ and $\tilde{p}(A)$, the estimator of IS is $\hat{l}_{\mathrm{IS}} = (1/\tilde{c}_1)\sum_{j=1}^{\tilde{c}_1}\{\mathbf{1}_{\{j\in\phi_1\}}\cdot p(\omega_{T_{\mathrm{In}\{1,j\}}})/\tilde{p}(\omega_{T_{\mathrm{In}\{1,j\}}})\}$, where $\tilde{c}_1$ and ϕ_1 denote the number of packets and the set of IDs of lost packets. The change-of-measure $p(\omega_{T_{\mathrm{In}\{1,j\}}})/\tilde{p}(\omega_{T_{\mathrm{In}\{1,j\}}})$ is analytically derived from a traffic model in model-based IS.

IV. Model-less Approach

Our goal is to accurately estimate a packet loss rate through a simulation by the change-of-measure based on traffic trace of $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ in a real network without assuming any traffic model. To accomplish this goal, we rewrite the estimator of model-based IS as $\hat{l}_{\mathrm{IS}} = (1/\tilde{c}_1)\int_0^\infty d_1(t)\cdot p(x_{1,t}\cap y_{1,t})/\tilde{p}(x_{1,t}\cap y_{1,t})\mathrm{d}t$ using the notation shown in Section II, where $A\cap B$ is a event that paths A and B occurs simultaneously. There is no way to strictly obtain $p(x_{1,t}\cap y_{1,t})/\tilde{p}(x_{1,t}\cap y_{1,t})$ except to analytically calculate assuming a traffic model like model based IS.

In the model-less approach, in order to obtain the change-of-measure from traffic trace of $x_i(t)$ and $y_i(t)$, we discretize $x_i(t)$ and $y_i(t)$ into Δ as $X_i(n) = \int_{(n-1)\Delta}^{n\Delta} x_i(t)\mathrm{d}t$ and $Y_i(n) = \int_{(n-1)\Delta}^{n\Delta} y_i(t)\mathrm{d}t\,(n\in\mathbb{N})$. In addition, $d_i(t)$ in a simulation is similarly discretized as $\tilde{D}_i(n) = \int_{(n-1)\Delta}^{n\Delta} d_i(t)\mathrm{d}t$.

Using these descritized traffic, we estimate the packet loss rate by

$$\hat{l}_i = \frac{1}{\tilde{c}_i}\sum_{N=1}^{\infty}\tilde{D}_i(N)\frac{P(\boldsymbol{X}_{N,k}\cap\boldsymbol{Y}_{N,k})}{\tilde{P}(\boldsymbol{X}_{N,k}\cap\boldsymbol{Y}_{N,k})}, \qquad (1)$$

where $P(A)$ and $\tilde{P}(A)$ denote the probabilities that event A occurs in the traffic trace and the simulation, respectively. $\boldsymbol{X}_{N,k} = \{X_{i,N,k}\}_{i\in\mathbb{N}}$ and $\boldsymbol{Y}_{N,k} = \{Y_{i,N,k}\}_{i\in\mathbb{N}}$ are discritized version of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. $X_{i,N,k}$ and $Y_{i,N,k}$ are defined as $\{X_i(n)\}_{N-k<n\leqq N}$ and $\{Y_i(n)\}_{N-k<n\leqq N}$, respectively. k is a parameter and we assume that the effect of $\{X_i(n)\}_{n\leq N-k}$ and $\{Y_i(n)\}_{n\leq N-k}$ on packet losses at time N is negligible. When we assume a single router and a single flow, in the limit as $\Delta\to 0$ and $k\to N$, our estimator converges to that of model based IS.

Our approach can estimate a packet loss rate without the restrictions of the conventional IS. Because of the discretization, $P(\boldsymbol{X}_{N,k}\cap\boldsymbol{Y}_{N,k})$ directly obtained from arbitrary traffic trace of $\boldsymbol{x}(t)$ without modeling the traffic. In our method, we perform a simulation using the simple Poisson traffic model irrespective of the traffic trace. Therefore, it is easy to analytically obtain $\tilde{P}(\boldsymbol{X}_{N,k}\cap\boldsymbol{Y}_{N,k})$. By expressing the estimator by input and output traffic instead of a queue length process, (1) is applicable for multiple flows on a network with complicated topology.

V. Experiments

As a first step in the development, we investigate the case when the packet loss rate of an MMPP/M/1/K system is estimated from an M/M/1/K simulation. MMPP of the arrival process in the original system has 2-states with traffic intensity 100 [packet/s] and 339 [packet/s]. The transition rate of each state is 1 times per second. For both of the original system and the simulation, the mean service time is 0.001 [s] and the queue size is 10 [packet]. The simulation time is 2000 [s], $\Delta = 0.025$ [s], and $k = 2$. In these systems, since the packet arrivals and a service time are independent, it is assumed that $P(Y_{1,N,k}|X_{1,N,k})/\tilde{P}(Y_{1,N,k}|X_{1,N,k}) = 1$. Therefore, the change-of-measure can be expressed as $P(X_{1,N,k}\cap Y_{1,N,k})/\tilde{P}(X_{1,N,k}\cap Y_{1,N,k}) = P(X_{1,N,k})/\tilde{P}(X_{1,N,k})$.

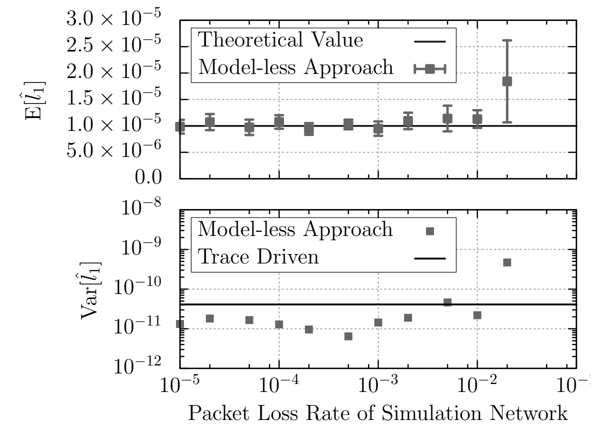

Fig. 2. Comparison between the estimators of the model-less approach and the trace-driven simulation.

By changing the loss rate in the simulation through the traffic intensity, we performed the simulation 30 times for each loss rate. Then the average of the estimator with 95% confidence intervals and the variance of the estimator were calculated. The result is compared with the trace-driven simulation in Fig. 2. According to the figure, we can find the region in which the model-less approach can estimate the packet loss rate of the original system. Additionally, we can confirm that the variances of the estimators are about 1/3 in the region, compared with the estimator by the trace-driven simulation.

VI. Conclusions and Future Directions

We proposed the model-less approach to accurately estimate a packet loss rate through simulation with traffic trace without traffic models. Though our approach is applicable to arbitrary trace on arbitrary network with a complicated topology, we only confirmed its effectiveness on a single router with a single flow in this paper. We will verify the applicability of our approach to the various trace on various networks in our future works.

References

[1] J. Zhang *et al.*, "A Survey of Network Traffic Generation," in *Proc. of CCT 2015*.

[2] N. Kobayasahi *et al.*, "On-line Estimation by Importance Sampling for the Tail Probability of FIFO Queue Length," in *Proc. of RESIM 2014*.

[3] I. C. Paschalidis *et al.*, "Importance Sampling for the Estimation of Buffer Overflow Probabilities via Trace-driven Simulations," *IEEE/ACM Transactions on Networking*, vol. 12, no. 5, 2004.

[4] P. E. Heegaard *et al.*, "Application of Rare Event Techniques to Trace Driven Simulation," in *Proc. of WSC 2005*.