Model-less Approach for an Accurate Packet Loss Simulation

1st Kohei Watabe
Graduate School of Engineering, Nagaoka University of Technology
Nagaoka, Niigata, Japan
k_watabe@vos.nagaokaut.ac.jp

2nd Masahiro Terauchi
Graduate School of Engineering, Nagaoka University of Technology
Nagaoka, Niigata, Japan

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

***Abstract*—In network evaluation through simulations, accurately modeling traffic of real networks is difficult. Even if accurate traffic modeling is achieved, it is also difficult to accurately estimate a rate of rare packet loss events. For accurate estimations of rare events, Importance Sampling (IS) based on the change-of-measure technique using traffic models has been investigated. However, these studies are inapplicable for traffic traces of real networks since the applicable traffic models are extremely limited. In this paper, we propose a model-less approach to accurately estimate a packet loss rate through a simulation without directly modeling traffic. The change-of-measure is achieved based on traffic traces of networks in our model-less approach. We evaluated the applicability of the model-less approach on a G/M/1/K system with a traffic trace of a real network and confirmed that the model-less approach achieves up to 145 times accurate than normal a trace-driven Monte Carlo (MC) simulation.**

***Index Terms*—Network Simulation, Packet Loss Rate, Importance Sampling, Rare Event Simulation**

I. Introduction

Network simulation is a fundamental technique to evaluate Quality of Service (QoS) on networks. Especially for large-scale networks, a low-cost evaluation can be achieved by a simulation technique before Internet Service Providers (ISPs) or other companies actually build their networks. Moreover, QoS evaluation through network simulations provides helpful information when we develop emerging applications/protocols on networks.

When we evaluate QoS of networks through simulations, accurately modeling traffic of real networks is an important but difficult task. Many traffic models that reflect various characteristics of the traffic measured in real networks have been proposed by prior works (see [2] and the references therein). The most fundamental one is a Poisson traffic model. It was reported that traffic of real networks has various characteristics, including long-range dependency [3], self-similarity [4], periodicity [5], which are different from the Poisson traffic model. Therefore, we need to choose the best traffic model from many traffic models, according to a situation that a simulation reproduces. However, due to the variety of traffic models, it is difficult to select an appropriate traffic model and tune its parameters for the traffic of the real network (The double arrows in Fig. 1).

This work was partly supported by JSPS KAKENHI Grant Number JP17K00008 and JP18K18035. An earlier version of this paper was presented at the 26th IEEE International Conference on Network Protocols (ICNP 2018) Poster session [1].

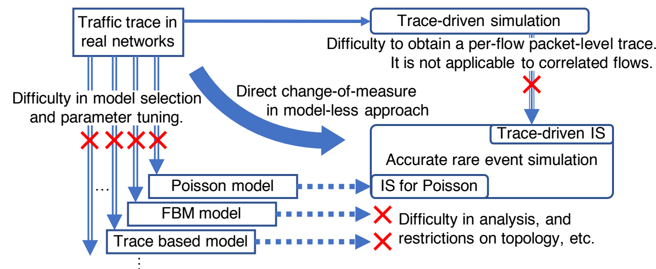

Fig. 1. The conventional approach and the model-less approaches for an accurate estimation with a traffic trace in a real network.

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 emerging applications with strict performance needs, ITU-T Recommendation Y.1547 [6] defines a QoS class with an upper bound 10^{-5} on a packet loss rate. Packet loss events are rare events in networks with such a QoS class, and it is a very difficult task to estimate the packet loss rate through a simulation.

For accurate estimations of rare events, Importance Sampling (IS) in which a simulation of a network where the events occur more frequently is performed has been used [7]–[9]. In the IS method for networks, the result of the original network is obtained by change-of-measure from the result of the network in the simulation. In most of the studies regarding the IS method in networks, the change-of-measure is analytically calculated based on a traffic model (model-based IS). They are inapplicable for real network traffic since the applicable traffic models, network topology, etc. are extremely limited, e.g. Poisson traffic model on a single router, etc. (The dotted arrows in Fig. 1).

Although there are a few works regarding IS in trace-driven simulation without traffic models [10], [11] (The triple arrows in Fig. 1), they require a per-flow packet-level traffic trace in which packet arrivals of each flow are recorded. In trace-driven IS, traffic traces of multiple flows that are measured in a real network is used, and the phases of the flow traffic traces are shifted in order to cause congestion. Because of its structure, change-of-measure cannot be obtained analytically unless they are uncorrelated in a time-space. Unfortunately, per-flow packet-level traffic traces are hardly disclosed in public. Furthermore, since these traces in real networks are definitely correlated, trace-driven IS is not suitable for traffic traces in real networks. In addition, to the best of our knowledge, no study on trace-driven IS has been published after the studies [10], [11], 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 (The bold arrow in Fig. 1). The model-less approach provides change-of-measure with a traffic trace of a network, including real networks. Though traffic of a simulation actually performed in the model-less approach is totally different from a traffic trace, an appropriate estimation can be achieved. Despite the packet loss rate in the original network to be evaluated naturally does not correspond to that of the simulation, the estimation result of the loss rate for a traffic trace is obtained by a change-of-measure technique. Though change-of-measure in the model-less approach is based on model-based IS, it utilizes a frequency distribution of a discretized traffic trace instead of a traffic model. By a direct change-of-measure with a traffic trace, we can avoid the difficulty of model selection (The double arrows in Fig. 1), the difficulty in analysis, and the restrictions on topology (The dotted arrows in Fig. 1). The model-less approach has potential applicability for queueing networks with complex topologies though we evaluate the model-less approach for a G/M/1/K as a first step in the development in this paper.

The main contributions of this paper are as follows.

- We construct a theoretical framework to derive change-of-measure without direct modeling of network traffic. Change-of-measure is achieved with a traffic trace of a network, and there is no assumption regarding traffic characteristics.
- We design an algorithm to generate traffic of a simulation actually performed. By the algorithm, the simulation traffic in which rare events in a traffic trace frequently occur is generated from a traffic trace, and an accurate estimation of a packet loss rate is achieved.
- Through simulations including a traffic trace of a real network, we confirm the effectiveness of the model-less approach on a G/M/1/K system.

The remainder of this paper is organized as follows. First, we summarize related work in Section II, and detail model-based IS in Section III. Next, we formulate the problem we tackle in Section IV. Section V explains the model-less approach, and Section VI evaluate the proposed method with Poisson traffic traces and a traffic trace in a real network. Lastly, in Section VII, we conclude the paper and mention future directions of the research.

II. Related Work

IS has been developed by many researchers in the literature of rare event simulation (see [12] and references therein). In the research field of queueing theory, many researchers have investigated how to achieve an accurate/fast estimation of statistics regarding rare events including buffer overflow events [13]. IS in queueing networks with complex topologies have also been tried, and Dupuis *et al.* investigated buffer overflows probability in stable open Jackson networks [14]. Mahdipour *et al.* evaluated the performance of IS in Jackson networks [8]. Some researchers have tackled optimization problems of estimation accuracy. Kim *et al.* proposed a non-parametric adaptive IS using a kernel function estimation method to explore the optimal condition of a simulation [15]. de Boer *et al.* also proposed an optimization method based on cross-entropy, and they confirmed the effectiveness of the method in Jackson networks [16].

There are a few works that are applicable for traffic traces of real networks while many works regarding IS techniques have been applied for markovian queueing networks. Paschalidis *et al.* proposed a trace-driven IS method in which phase of multiple traffic traces are synchronized to lead congestion in a system [10]. By using change-of-measure, the result in a system with synchronized traffic traces are converted to a result in a system where the phase of traffic traces are randomized. Reference [11] also proposed a simulation technique with a similar idea. However, these studies structurally inapplicable for systems with correlated flows, and traffic traces of parallel flows are generally correlated. Additionally, traffic traces are not always measured as per-flow packet-level traces. In the literature of general IS that does not focus on queueing systems, an IS technique for a system in which part of the system is not directly known was proposed [17]. Regrettably, to the best of our knowledge, the studies that apply the technique to the queueing systems is not reported.

III. Model-based IS

In model-based IS, a simulation of a network where the target events occur more frequently than the original network is performed, and the result of the simulation is converted to the result of the original network. When we estimate a packet loss rate of a network by model-based IS, we intentionally increase the traffic intensity of the queueing model that expresses the network in order to cause loss events frequently. The loss rate of the original network can be obtained by the change-of-measure technique.

When we estimate the loss rate on a single router with a single queue, the estimator of model-based IS for the packet loss rate can be given by

$$\hat{l}_{\mathrm{IS}} = \frac{1}{\tilde{c}} \sum_{j=1}^{\tilde{c}} \left\{ \mathbf{1}_{\{j \in \phi\}} \frac{p(\tilde{\omega}_j)}{\tilde{p}(\tilde{\omega}_j)} \right\}, \tag{1}$$

where $\tilde{c}$ and ϕ represent the number of packets in a simulation and a set $\{1, 2, \ldots, \tilde{c}\}$ of packet IDs, respectively. $\tilde{\omega}_j$ denotes a path of a queue length process of the router until arrival

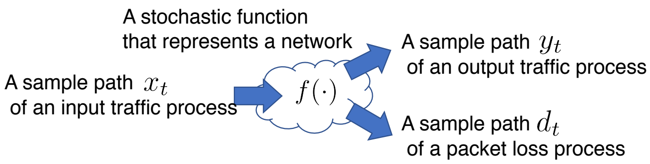

Fig. 2. A formulation in the model-less approach.

time s_j of jth packet in a simulation. Namely, $\tilde{\omega}_j$ is a sample path of a queue length process in the period $[0, s_j]$. $p(A)$ and $\tilde{p}(A)$ represent the probability density of an event A in the original network and the simulation network, respectively. The part $p(\tilde{\omega_j})/\tilde{p}(\tilde{\omega_j})$ of Eq. (1), the ratio of the probability density of a queue length process in the original network to that of the simulation network, provides change-of-measure. If a simulation network corresponds to an original network, i.e. the condition is not changed, the estimator shown in Eq. (1) is equal to the estimator of a simple Monte-Carlo simulation since $\tilde{p}(A) = p(A)$.

In model-based IS, it is needed that the change-of-measure $p(\tilde{\omega_j})/\tilde{p}(\tilde{\omega_j})$ is analytically derived from the traffic model of the network. Therefore, the traffic model should be known and probability density should be analytically solved by using it. However, as we mentioned above, it is difficult to select/tuning traffic models according to a traffic trace of a real network, and simple traffic models that can be analytically treated cannot represent characteristics of the traffic of the real network. Consequently, we cannot accurately estimate the packet loss rate for a network in which packet loss events are rare events, with model-based IS.

IV. Problem Formulation

We treat a network as a stochastic function of input/output traffic in the formulation. Our formulation has high generality and applicability for traffic in real networks since we do not assume any restriction of input/output traffic characteristics. In this section, we describe a formulation of a packet loss rate estimation through a simulation (see Fig. 2).

In this paper, we express input and output traffic processes of a network by point processes $x(t)$ and $y(t)$, respectively. $x(t)$ and $y(t)$ are defined as the sum of Dirac delta functions, $x(t) = \sum_j \delta(t - s_j)$ and $y(t) = \sum_j \delta(t - u_j)$, where s_j and u_j denote arrival and departure time of jth packet, respectively. Let x_t and y_t be sample paths of $x(t)$ and $y(t)$ on a period $[0, T]$, respectively. Additionally, we also define a packet loss process $d(t)$ as $d(t) = \sum_j \delta(t - v_j)$, where v_j denotes the time when jth packet losses. Let d_t be a sample path of $d(t)$ on a period $[0, T]$.

Packet loss events stochastically occur depending on a sample path x_t of an input traffic process. A sample path y_t of an output traffic process and a sample path d_t of a packet loss process of a network can be represented by a stochastic function $f(\cdot)$ as $(y_t, d_t) = f(x_t)$. $f(\cdot)$ is a function that expresses the behavior of the network. We assume that the value at time t on the sample path d_t mainly depends on the last τ period $[t - \tau, t)$ of traffic process, and the value at time t on d_t is approximately independent on x_t in the period $[0, t - \tau)$ for sufficiently large τ. Needless to say, the output traffic process $y(t)$ and the packet loss process $d(t)$ are stochastic processes, and there are two sources of the randomness of the two processes: stochastic functions $f(\cdot)$ and an input traffic process $x(t)$. A packet loss rate l is defined by

$$l = \frac{\int_0^T d_t \mathrm{d}t}{\int_0^T x_t \mathrm{d}t}.$$

We assume that the packet loss rate l is extremely low, and it is highly dependent on a sample path x_t of an input traffic process. Since the main cause of packet loss events that we assume is a buffer overflow, the dependency of a packet loss rate on a traffic process is obvious.

The problem we tackle in this paper is a simulation-based estimation of an average packet loss rate $\mathrm{E}[\,l\,]$ from a measurement of traffic trace x_t and a simulation of behavior $f(\cdot)$ of a network. We assume that behavior $f(\cdot)$ of a network can be accurately simulated by a simulation. Many packet-level simulators including ns-3 simulator [18] have been proposed in the literature on network simulations, and these outstanding simulators achieve accurate reproduction of network behavior. This assumption means that the following two equations hold:

$$p(y_t|x_t) = \tilde{p}(y_t|x_t), \qquad (2)$$
$$p(d_t|x_t) = \tilde{p}(d_t|x_t),$$

where $p(A|B)$ and $\tilde{p}(A|B)$ denote conditional probability density that an event A occurs under the condition that an event B occurs in an original network and a simulation network, respectively.

Since it is relatively hard to simulate traffic by a traffic model, as we mentioned above, we try to utilize a traffic trace of a real network in a simulation. Of course, it is easy to perform a trace-driven Monte Carlo (MC) simulation by directly input the traffic trace into a network simulator, i.e., we can obtain d_t by simulating $f(x_t)$. In this paper, however, we tackle to develop a method in which simulation traffic $\tilde{x}_t$ different from a traffic trace x_t is simulated as $(\tilde{y}_t, \tilde{d}_t) = f(\tilde{x}_t)$, thereby accurately estimating the packet loss rate. An output traffic process $\tilde{y}_t$ and a packet loss process $\tilde{d}_t$ are obtained by the simulation. Though $\tilde{d}_t$ in the simulation differs from d_t, the packet loss rate l is appropriately derived by a change-of-measure technique.

V. Model-less Approach

Our goal is to accurately estimate a packet loss rate by an appropriate change-of-measure through a network simulation. Aiming to achieve the goal, we propose the model-less approach whose steps is shown in Fig. 3. In the model-less approach, the change-of-measure is achieved based on a traffic trace x_t that is measured in a limited measurement period $[0, T]$ on a real network. Though traffic in a simulation is completely different from a traffic trace x_t in a real network, the packet loss rate should be estimated without bias by the

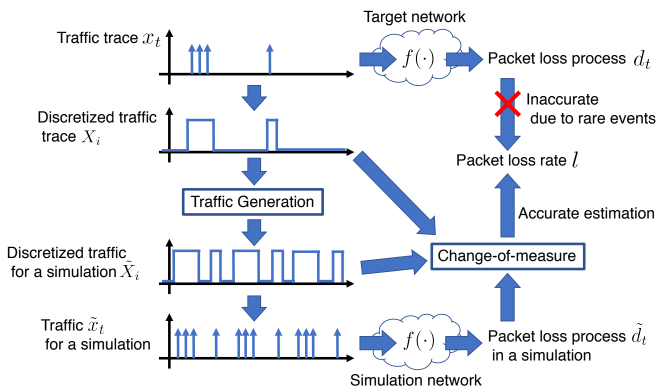

Fig. 3. Steps for an accurate estimation in the model-less approach.

change-of-measure. In this section, we will explain a change-of-measure technique that achieves an unbiased and accurate estimation, and how to generate traffic $\tilde{x}_t$ in a simulation.

A. Change-of-measure Technique

The model-less approach avoids analytical derivation of change-of-measure, and enable an empirical derivation of change-of-measure from a traffic trace. In order to achieve empirical derivation of change-of-measure, we discretize a traffic process and decompose it into patterns. Occurrence frequency of each traffic pattern is counted, and change-of-measure is calculated from a frequency distribution of traffic patterns.

We design the proposed method to have high versatility by eliminating restrictions (e.g., characteristics of a traffic trace, network topology, and so on) while acting as an extension of the conventional model-based IS. Using notations we defined in Section IV, we can rewrite the estimator shown in Eq. (1) of model-based IS as

$$\begin{aligned}\frac{1}{\tilde{c}}\sum_{j=1}^{\tilde{c}}\left\{\mathbf{1}_{\{j\in\phi\}}\frac{p(\tilde{\omega}_j)}{\tilde{p}(\tilde{\omega}_j)}\right\} &= \frac{1}{\tilde{c}}\int_0^T \tilde{d}_s\frac{p(\tilde{x}_t^s\wedge\tilde{y}_t^s)}{\tilde{p}(\tilde{x}_t^s\wedge\tilde{y}_t^s)}\mathrm{d}s\\ &= \frac{1}{\tilde{c}}\int_0^T \tilde{d}_s\frac{p(\tilde{x}_t^s)p(\tilde{y}_t^s|\tilde{x}_t^s)}{\tilde{p}(\tilde{x}_t^s)\tilde{p}(\tilde{y}_t^s|\tilde{x}_t^s)}\mathrm{d}s\\ &= \frac{1}{\tilde{c}}\int_0^T \tilde{d}_s\frac{p(\tilde{x}_t^s)}{\tilde{p}(\tilde{x}_t^s)}\mathrm{d}s, \qquad (3)\end{aligned}$$

where $\tilde{x}_t^s$ and $\tilde{y}_t^s$ denote sample paths of input and output traffic processes until time s in a simulation, respectively. Note that the operator $\wedge$ denotes logical conjunction. The last equality follows from Eq. (2). The change-of-measure is achieved by the part $p(\tilde{x}_t^s)/\tilde{p}(\tilde{x}_t^s)$ of Eq. (3). However, when we derive $p(\tilde{x}_t^s)/\tilde{p}(\tilde{x}_t^s)$, it is required to analytically calculate $p(\tilde{x}_t^s)/\tilde{p}(\tilde{x}_t^s)$ assuming traffic model since probability density $p(\tilde{x}_t^s)$ and $\tilde{p}(\tilde{x}_t^s)$ of an event $\tilde{x}_t^s$ are defined on a space $\mathbb{R}^{\tilde{c}}$.

According to the assumption of independence of d_t on x_t in $[0, t-\tau)$, we can rewrite Eq. (3) as

$$\frac{1}{\tilde{c}}\int_0^T \tilde{d}_s\frac{p(\tilde{x}_t^{s,\tau})}{\tilde{p}(\tilde{x}_t^{s,\tau})}\mathrm{d}s, \qquad (4)$$

where $\tilde{x}_t^{s,\tau}$ denotes a part of the sample path $\tilde{x}_t$ in the last τ period $[s-\tau, s)$.

In the model-less approach, in order to achieve empirical derivation of change-of-measure based on a traffic trace, we replace $\tilde{x}_t$ and $\tilde{d}_t$ by sample paths $\tilde{X}_i$ and $\tilde{D}_i$ of discretized processes, respectively. $\tilde{x}_t$ and $\tilde{d}_t$ are divided into discrete slots with interval Δ.

$$\tilde{X}_i = \int_{(n-1)\Delta}^{n\Delta}\tilde{x}_t\mathrm{d}t, \quad \tilde{D}_i = \int_{(n-1)\Delta}^{n\Delta}\tilde{d}_t\mathrm{d}t.$$

x_t is also discretized as X_i. Similarly, we define $\tilde{X}_i^{n,k} = \{\tilde{X}_i\}_{n-k<i\le n}$ as a discretized version of $\tilde{x}_t^{s,\tau}$. By defining $\tilde{X}_i^{n,k}$ as a discretized process, $p(\tilde{x}_t^{s,\tau})$ and $\tilde{p}(\tilde{x}_t^{s,\tau})$ are also replaced by $P(\tilde{X}_i^{n,k})$ and $\tilde{P}(\tilde{X}_i^{n,k})$, respectively. $P(A)$ and $\tilde{P}(A)$ denote probability that an event A occurs in an original network and a simulation network, respectively. Consequently, the estimator is rewritten as

$$\frac{1}{\tilde{c}}\sum_{n=1}^{T/\Delta}\tilde{D}_i\frac{P(\tilde{X}_i^{n,k})}{\tilde{P}(\tilde{X}_i^{n,k})}. \qquad (5)$$

The part $P(\tilde{X}_i^{n,k})/\tilde{P}(\tilde{X}_i^{n,k})$ in Eq. (5) corresponds the change-of-measure, and the probability $P(\tilde{X}_i^{n,k})$ and $\tilde{P}(\tilde{X}_i^{n,k})$ of an event $\tilde{X}_i^{n,k}$ are defined on a space $\mathbb{N}^k$. It is important to limit a period of the sample path used in change-of-measure in Eq. (4). The limitation enable us to calculate the probability $P(\tilde{X}_i^{n,k})$ and $\tilde{P}(\tilde{X}_i^{n,k})$ by counting frequency of a traffic pattern $\tilde{X}_i^{n,k}$ in X_i and $\tilde{X}_i$.

B. Traffic Generation in Simulations

In the model-less approach, traffic patterns that are observed in a traffic trace of an original network are listed, and input traffic $\tilde{x}_t$ is generated for a simulation so that the frequency of all patterns are the same. Traffic patterns referred to here are $X_i^{n,k}$ that is k consecutive slots in a discretized traffic trace X_i. We define Ω as the set of all traffic patterns appearing in X_i without duplicates. By making the frequency of all patterns are the same, the frequency of the patterns that cause loss events increases since we assume that the loss events are rare. It is not a good policy to break the uniformity of the frequency of the patterns. If we raise the frequency of specific patterns, the frequency of the other patterns has to be decreased thereby being difficult to calculate the probability $P(\tilde{X}_i^{n,k})$ and $\tilde{P}(\tilde{X}_i^{n,k})$ from the low frequency of patterns.

The algorithm to generate input traffic $\tilde{x}_t$ with the uniform frequency of all patterns is shown in Algorithm 1. The input parameters of the algorithm are the following four parameters: discretized traffic trace X_i, slot interval Δ, the number k of a consecutive slot in a traffic pattern, and a simulation period T_{sim}. The output of the algorithm is discretized traffic $\tilde{X}_i$ for a simulation. Parameters are initialized on Line 1–3, and the number N_{sim} of slots in discretized traffic $\tilde{X}_i$ is calculated in Line 4. Note that $\lceil\cdot\rceil$ denotes a ceiling function. The loop on Line 5–10 actually generates $\tilde{X}_i$. On Line 6, among the traffic patterns that can be joined to current $\tilde{X}_i$, a traffic pattern having the lowest frequency in simulation traffic $\tilde{X}_i$

Algorithm 1 Traffic generation algorithm for simulations

Require: $X_i, \Delta, k, T_{\mathrm{sim}}$
Ensure: $\tilde{X}_i$
1: $\tilde{X}_i \leftarrow [0],\ \kappa \leftarrow k-1$
2: **for all** $\chi \in \Omega$ **do**
3: $\quad h(\chi) \leftarrow 0$
4: $N_{\mathrm{sim}} \leftarrow \lceil T_{\mathrm{sim}}/\Delta \rceil$
5: **while** length of $\tilde{X}_i < N_{\mathrm{sim}}$ **do**
6: $\quad \chi' \leftarrow$ Randomly sample from $\underset{I_\kappa(\tilde{X}_i,\chi)=1}{\arg\min}\ h(\chi)$
7: $\quad$ **if** $\chi' = \emptyset$ **then** $\kappa \leftarrow \kappa - 1$
8: $\quad$ **else**
9: $\quad\quad [\chi'[\kappa+1], \ldots, \chi'[k]]$ is joined to $\tilde{X}_i$
10: $\quad\quad h(\chi') \leftarrow h(\chi') + 1$
11: $\quad\quad \kappa \leftarrow k-1$

is randomly chosen. $h(\chi)$ denotes frequency of a pattern χ in $\tilde{X}_i$. $I_\kappa(\tilde{X}_i, \chi)$ is an indicator function which equals 1 when the last κ slots of $\tilde{X}_i$ correspond to the first κ slots of χ and 0 otherwise. If there is no χ satisfying the condition of the argument of the minimum, κ is reduced thereby relaxing the condition that the last κ slots of $\tilde{X}_i$ correspond to the first κ slots of χ (Line 7). When χ' is not empty, the selected pattern χ' is joined to $\tilde{X}_i$ with an overlap of κ slots (Line 9). After the joining, frequency $h(\chi)$ of traffic patterns in $\tilde{X}_i$ is updated on Line 10, and κ is initialized on Line 11.

Lastly, the discretized traffic $\tilde{X}_i$ that is obtained by Algorithm 1 is converted to $\tilde{x}(t)$. The value of each time slot of the discretized traffic $\tilde{X}_i$ represents the number of arrival packets in the time slot. Hence, we generate simulation traffic $\tilde{x}(t)$ by assuming the arrival time of packets s_j is uniformly distributed in a time slot.

VI. Evaluations

In this paper, in order to evaluate the effectiveness of the model-less approach, a packet loss rate is estimated in a simple queueing model by the model-less approach and trace-driven MC simulations. In the evaluations, we prepared the following two traffic traces as input traffic: Poisson traffic trace and a traffic trace of a real network in Widely Integrated Distributed Environment (WIDE) project [19]. These input traffic are used as an arrival process of G/M/1/K (i.e., M/M/1/K for Poisson traffic trace). We compare the model-less approach and trace-driven MC simulations in which the traffic trace is directly input into a G/M/1/K queueing system. Note that the trace-driven MC simulation in the evaluation is not a trace-driven IS simulation we introduced in the introduction.

A. Poisson Traffic Trace

Firstly, we evaluate the most fundamental case where input traffic follows the Poisson traffic model. However, the model-less approach does not use the knowledge that the input traffic follows the Poisson traffic model. By evaluating an M/M/1/K queueing system where the true value of the packet loss rate

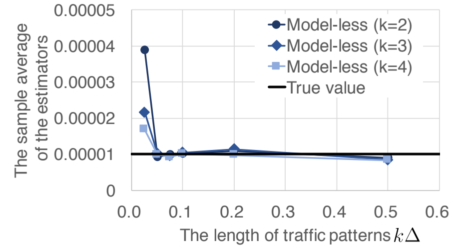

Fig. 4. The sample average of the estimator of the packet loss rate in an M/M/1/K queueing system.

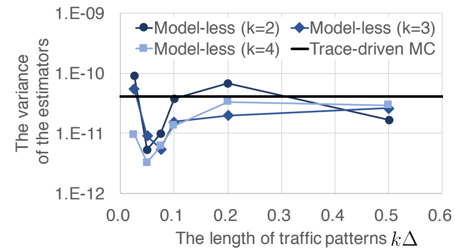

Fig. 5. The variance of the estimator of the packet loss rate in an M/M/1/K queueing system.

is known, we can confirm whether the estimator of the model-less approach has a bias or not.

The parameters of an M/M/1/K queueing system was tuned so that a true value of the packet loss rate is 10^{-5}. 10^{-5} is also a criterion defined in the most severe QoS class in ITU-T Recommendation Y.1547 [6]. The arrival rate and the service rate are set to $\lambda = 329.1$ [packet/s] and $\mu = 1000$ [packet/s], respectively. The queue length K is set to 10 [packet]. We perform simulations 30 times, and the period for each simulation is 2000 [s].

The results of the sample average of the estimators of the packet loss rate are shown in Fig. 4. The horizontal and the vertical axis represent the length of a traffic pattern $k\Delta$ and sample average $\mathrm{E}[\,\hat{l}\,]$ of the estimators. The black line at 1.0×10^{-5} represents the true value of the packet loss rate in the M/M/1/K queueing system. According to the figure, we can confirm that the model-less approach estimates the packet loss rate without bias when $k\Delta \geq 5 \times 10^{-2}$.

Figure 5 shows the variance of the estimators of the model-less approach and the trace-driven MC simulation. The horizontal and the vertical axis represent the length $k\Delta$ of a traffic pattern and variance $\mathrm{Var}[\,\hat{l}\,]$ of the estimators. The black line at 4.0×10^{-11} represents not true value but the variance of the estimators of the trace-driven MC simulation. According to the figure, most of the results of the model-less approach is lower than that of the trace-driven MC simulation when the length of a traffic pattern $k\Delta \geq 5 \times 10^{-2}$ [s]. Especially, the variance of the estimators of the model-less approach is about 1/12 of that of the trace-driven MC simulation when $k = 4$ and $k\Delta = 5 \times 10^{-2}$ [s]. Since the estimators of the model-less approach are unbiased for $k\Delta \geq 5 \times 10^{-2}$ [s] in Fig. 4, we can confirm that the model-less approach provides an accurate estimator of the packet loss rate.

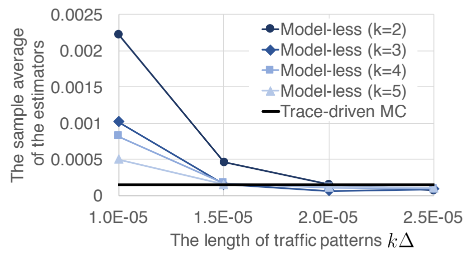

Fig. 6. The sample average of the estimator of the packet loss rate in a G/M/1/K queueing system with the traffic trace of the real network.

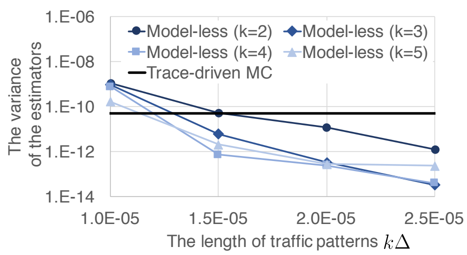

Fig. 7. The variance of the estimator of the packet loss rate in a G/M/1/K queueing system with the traffic trace of the real network.

B. *Traffic Trace of a Real Network*

In order to verify the applicability of the model-less approach to traffic traces of real networks, we evaluate the model-less approach with a traffic trace opened by MAWI (Measurement and Analysis on the WIDE Internet) Working Group in WIDE project [19]. The traffic trace we used was measured at samplepoint-F on September 8, 2018, and we extracted the arrival time of packets from the traffic trace. It is difficult to analytically derive the true value of the packet loss rate of a G/M/1/K queueing system with the traffic trace of the real network. For that reason, we tune the service rate of the G/M/1/K queueing system so that the packet loss rate of the trace-driven MC simulation is approximately 10^{-4}, thereby setting $\mu = 10^6$ [packet/s] as a service rate. The queue length K is set to 10 [packet]. We perform simulations 10 times, and a period for each simulation is 100 [s].

The estimation results of the model-less approach and the trace-driven MC simulation are shown in Figs. 6 and 7. According to the figures, we can confirm that an unbiased and accurate estimation is achieved if k and $k\Delta$ are sufficiently large. In particular, the variance of the estimator of the model-less approach is 1/145 of that of the trace-driven MC simulation when $k = 8$ and $k\Delta = 2.0 \times 10^{-5}$ [s].

VII. Conclusions and Future Works

In this paper, we proposed the model-less approach to accurately estimate a packet loss rate through a simulation without directly modeling traffic. The model-less approach provides a change-of-measure technique based on model-based IS with a frequency distribution of discretized traffic patterns. We evaluated the model-less approach in G/M/1/K with a Poisson traffic trace and the traffic trace in the real network, and confirmed that the model-less approach can achieve an unbiased and accurate estimation. In the most effective cases in the evaluation, the accuracy of the model-less approach is 145 times accurate than that of the normal trace-driven MC simulation.

In future works, we will tackle the optimization problem of the parameters of the model-less approach. The length of the patterns and the interval of discrete slots should be specified depending on characteristics of original networks and length of traffic traces.

References

[1] M. Terauchi, K. Watabe, and K. Nakagawa, "Model-Less Approach of Network Traffic for Accurate Packet Loss Simulations," in *Proceedings of the 26th IEEE International Conference on Network Protocols (ICNP 2018) Poster session*, Cambridge, UK, 2018, pp. 251–252.

[2] J. Zhang, J. Tang, X. Zhang, W. Ouyang, and D. Wang, "A Survey of Network Traffic Generation," in *Proceedings of the 3rd International Conference on Cyberspace Technology (CCT 2015)*, Beijing, China, 2015.

[3] T. Karagiannis, M. Molle, and M. Faloutsos, "Long-Range Dependence: Ten Years of Internet Traffic Modeling," *IEEE Internet Computing*, vol. 8, no. 5, pp. 57–64, 2004.

[4] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation*. John Wiley & Sons, 2000.

[5] S. Floyd and V. Jacobson, "The Synchronization of Periodic Routing Messages," *IEEE/ACM Transactions on Networking*, vol. 2, no. 2, pp. 122–136, 2002.

[6] "Network Performance Objectives for IP-based Services," *ITU-T Recommendation Y.1541*, 2011.

[7] J. Blanchet and H. Lam, "State-dependent Importance Sampling for Rare-event Simulation: An Overview and Recent Advances," *Surveys in Operations Research and Management Science*, vol. 17, no. 1, pp. 38–59, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.sorms.2011.09.002

[8] E. B. Mahdipour, A. M. Rahmani, and S. Setayeshi, "Performance Evaluation of an Importance Sampling Technique in a Jackson Network," *International Journal of Systems Science*, vol. 45, no. 3, pp. 373–383, 2014.

[9] I. Lokshina, "Study on Estimating Probabilities of Buffer Overflow in High-speed Communication Networks," *Telecommunication Systems*, vol. 62, no. 2, pp. 289–302, 2016.

[10] I. C. Paschalidis and S. Vassilaras, "Importance Sampling for the Estimation of Buffer Overflow Probabilities via Trace-driven Simulations," *IEEE/ACM Transactions on Networking*, vol. 12, no. 5, pp. 907–919, 2004.

[11] P. E. Heegaard, B. E. Helvik, and R. O. Andreassen, "Application of Rare Event Techniques to Trace Driven Simulation," in *Proceedings of the Winter Simulation Conference (WSC 2005)*, Orlando, FL, USA, 2005, pp. 509–518.

[12] J. Morio, M. Balesdent, D. Jacquemart, and C. Vergé, "A Survey of Rare Event Simulation Methods for Static Input-output Models," *Simulation Modelling Practice and Theory*, vol. 49, pp. 287–304, 2014.

[13] D. Miretskiy, *Queueing Networks: Rare Events and Fast Simulations*, 2009. [Online]. Available: http://doc.utwente.nl/68376/

[14] P. Dupuis and H. Wang, "Importance Sampling for Jackson Networks," *Queueing Systems*, vol. 62, no. 1-2, pp. 113–157, 2009.

[15] Y. B. Kim, D. S. Roh, and M. Y. Lee, "Nonparametric Adaptive Importance Sampling for Rare Event Simulation," in *Proceedings of the 32nd Conference on Winter Simulation (WSC 2000)*, Orlando, FL, USA, 2000, pp. 767–772.

[16] P. T. de Boer, D. P. Kroese, and R. Y. Rubinstein, "A Fast Cross-Entropy Method for Estimating Buffer Overflows in Queueing Networks," *Management Science*, vol. 50, no. 7, pp. 883–895, 2004.

[17] H. Lam, "Efficient importance Sampling Under Partial Information," in *Proceedings of the 2012 Winter Simulation Conference*, Berlin, Germany, 2012, pp. 454–465.

[18] T. R. Henderson, M. Lacage, G. F. Riley, G. Dowell, and J. B. Kopena, "Network Simulations with the ns-3 Simulator," in *Proceedings of ACM SIGCOMM 2008*, Seattle, WA, USA, 2008, p. 527.

[19] "WIDE MAWI WorkingGroup." [Online]. Available: http://mawi.wide.ad.jp/