An Adaptive Technique for Estimating Traffic Intensities for a Network in Random Markovian Routing

DEBANJAN BHATTACHARJEE, YAN ZHUANG

Department of Mathematics, Utah Valley University, Orem, UT, USA
Department of Mathematics and Statistics, Connecticut College, New London, Connecticut, USA

Abstract: We propose an efficient method for estimating a vector of parameters that characterize the node-to-node traffic intensities in a network. The estimation procedure is free from any initial estimate which is often difficult to choose for iterative procedures that occur in maximum likelihood estimation or method of moments. The procedure discussed here updates the information from a continuous chain of data until a stopping variable decides the final sample size and the estimate. More importantly, it considers the cost of logistics in sampling. We discuss important desirable properties of the estimation-procedure under the assumption of Poisson distribution. Data analyses reveal a highly accurate estimation and demonstrate a fast convergence.

Key-Words: Markovian-Routing, Network-Tomography, Poisson Thinning, Risk Efficiency, Sequential-Estimation

1 Introduction

The performance of a network, comprising node to node traffic, is often characterized by the delay in the system. In a network of c *source-destination (SD)* pairs, the quantum of signals passing from one node to another node is usually modeled by simple probability distributions such as Poisson distribution, characterized by the parameter-vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_c)$. The traffic-intensity in the links connecting these SD pairs is an important quantity to estimate. Under the Markovian routing, proposed in [1], the traffic intensities on the links may be characterized by a vector of linear combinations of $\lambda_1, \lambda_2, ..., \lambda_c$. Hence estimating such a vector of parameters is of great importance. This idea of Network Tomography [1] has received lots of attention recently in understanding the link performance in ad-hoc networks to VANET [2, 3, 4, 5, 6], and other related areas. Most recently, evaluation of software open-source IMS network under the load of SIP messages was addressed by [7] and identifying and correcting the network weaknesses by incorporating bandwidth measurements was discussed by [8].

The existing challenges in estimating the vector of parameters are quite non-trivial. Most of the literature, including [1], dealt with maximum likelihood estimation (MLE). Another method of estimation that has been adopted is based on method of moments also popularly know as EM (expectation-maximization) algorithm. Moreover, the system of linear equations to solve the parameters constitute a LININPOS (LINear INverse POSitive) problem [9, 10]. All these methods face substantial numerical difficulties. The convergence of the parameters, much of which is based on normal approximations, also depends on the initial choice in the iterative processes. We strive to circumvent these hindrances by proposing a simpler

and theoretically sound procedure.

It is very realistic to assume that the traffic data are obtained sequentially instead of several values coming in at one time [11]. In such a situation, where data arrive sequentially, an adaptive procedure, based on a certain stopping variable, may be more pragmatic to adopt. Such a procedure is not only efficient but may also save the number of required samples. If the data collection is costly, concerning time or money, an adaptive procedure can be implemented to control the cost of data collection. In this paper, we perform a minimum-risk point estimation of the average number of packets that pass through each link in a system of node-to-node traffic. We explore some attractive asymptotic properties for the proposed method of estimation. We also present empirical evidence through extensive simulation studies. In the literature, a comprehensive review of minimum-risk point estimation and sequential or adaptive sampling techniques are presented in [12]. A very recent work [13] presents the application of a sequential algorithm for anomaly detection in networks. There is a broad literature on theories and applications in sequential analysis, one is referred to [14], [15], and [12]. For more recent developments and applications utilizing sequential analysis, one is referred to [16], [17], [18], [19], and [20].

We summarize the main contributions of this paper as follows:

- The average number of messages at the links is extremely critical to estimate and the estimator is simple. The estimation process has a well-defined measure of accuracy, namely, *risk-efficiency*.

- A sequential method of recording data enables an estimation procedure to save the cost of estimation in terms of money, time, or other logistics,

without oversampling on average.

Section 2 of this article describes the basic probability distribution and some definitions similar to [1]. Section 3 conveys the estimation procedure and explores some appealing properties. Section 4 presents an extensive data analysis that supports the theoretical findings and the conjecture from Section 3. Finally, we conclude with some future directions.

2 The Probability Model

Consider c pairs of communicating nodes also called *source-destination (SD)* pairs and r directed links. Typically we face a situation where $c > r$. Following [1], we define a few necessary variables below. For $i = 1, 2, ..., r$ and $j = 1, 2, ..., c$, let

- $X_j \equiv$ the number of transmitted messages for the SD pair j.

- $Y_i^j \equiv$ the number of messages with address j passing through link i.

- $P_i^j \equiv$ probability that a message with address j passes through link i.

- $P_{ii'}^j \equiv$ probability that a message with address j passes through both link i and link i'.

- $Y_i \equiv$ total number of messages passing through the link i.

It is often assumed that the X_j's are distributed *independently* as Poisson with parameter λ_j for $j = 1, 2, ..., c$. Note that the quantity Y_i has its mean $\sum_j \lambda_j P_i^j$ for $i = 1, 2, ..., r$ and this is what we are interested to estimate. Moreover, because of the thinning property of Poisson distribution, the distribution of Y_i^j will be Poisson with mean $\lambda_j P_i^j$ for $i = 1, 2, ..., r$ and $j = 1, 2, ..., c$. We also note that for each $i = 1, 2, ..., r$

$$Y_i = \sum_{j=1}^{c} Y_i^j \sim \text{Poisson}\left(\sum_{j=1}^{c} \lambda_j P_i^j \right) \quad (1)$$

We also consider a routing matrix $\boldsymbol{A} = ((a_{ij}))_{r \times c}$ of probabilities of the propagation of the signal. To be more specific, we can say that a_{ij} is the conditional probability that a signal with SD-address $j = (j_1, j_2)$ passes through the link $i = (i_1, i_2)$ given that it leaves from the node i_1. Here, recall the basic nature of a Markov chain which says that future event would depend only on the most recent past. Hence, in our problem the Markovian nature of routing makes the probability a_{ij} independent of how the packet arrived at i_1. It rather depends on the final address j and only

the node from which it started going to j. For instance in Figure 1 from Section 4, the probability that a signal goes from node b to node d depends only on node b. It does not depend on the fact that the signal could arrive to b either from a or from c.

3 Estimation of Parameters

By using the theorem of total probability, we compute matrix $\boldsymbol{P} = ((P_i^j))_{r \times c}$ from the components of matrix $\boldsymbol{A}$. Here, P_i^j is $(i, j)^{th}$ component of the matrix $\boldsymbol{P}$. We assume that the matrix $\boldsymbol{A}$ is known to us so that $\boldsymbol{P}$ is completely specified. Then, by (1), we have,

$$E(\boldsymbol{Y}) = \boldsymbol{P}\boldsymbol{\lambda} \text{ and } Cov(\boldsymbol{Y}) = \boldsymbol{\Sigma} = ((\sigma_{ij}))_{r \times r} \quad (2)$$

where, $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_r)'$. Using definitions given in Section 2 and from equation (1) it can be shown that,

$$\sigma_{ii'} = \begin{cases} \sum_j \lambda_j P_{ii'}^j & \text{if} \quad i \neq i' \\ \sum_j \lambda_j P_i^j & \text{if} \quad i = i' \end{cases}$$

where, $i, i' = 1, 2, ..., r$ and $j = 1, 2, ..., c$.

The diagonal elements of the *variance-covariance* matrix $\boldsymbol{\Sigma}$ follows from (1). The off-diagonal elements of $\boldsymbol{\Sigma}$ or the covariance terms can be obtained by following the procedure given in [1] and for brevity, we will skip the details. Our main focus is on the estimation of the parameters which we discuss in the next section.

3.1 An Adaptive Procedure

We are going to adopt a *minimum-risk point estimation* procedure to estimate the mean vector $P\boldsymbol{\lambda}$. Having recorded $\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}, ..., \boldsymbol{Y}^{(n)}$, with, $\boldsymbol{Y}^{(k)} = (Y_1^{(k)}, Y_2^{(k)}, ..., Y_r^{(k)})'$ for $k = 1, 2, ..., n$, the sample mean vector, $\overline{\boldsymbol{Y}}_n = n^{-1} \sum_k \boldsymbol{Y}^{(k)}$ can be taken as an unbiased estimate of $P\boldsymbol{\lambda}$. We let the associated loss function to be of the form,

$$L_n(\boldsymbol{\lambda}, \overline{\boldsymbol{Y}}_n) = (\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})^T \boldsymbol{W}(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda}) + \gamma n \quad (3)$$

Here, $\gamma > 0$ is known, and $\boldsymbol{W}_{r \times r}$ is a known p.d. matrix. The term cn represents the cost of gathering n observations and $(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})^T \boldsymbol{W}(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})$ represents the loss due to the estimation of $\boldsymbol{P}\boldsymbol{\lambda}$. If n is small then the sampling cost γn will be small but $(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})^T \boldsymbol{W}(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})$ may be large. On the other hand if n is large the effects will be reversed. So the loss function in (3) achieves a trade-off between expenses due to sampling and estimation error.

Now using the fact that,

$$E_{\boldsymbol{\lambda}}\left[(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})(\overline{\boldsymbol{Y}}_n - \boldsymbol{P}\boldsymbol{\lambda})^T \right] = \tfrac{\boldsymbol{\Sigma}}{n},$$

the risk function associated with the loss function in (3) is given by,

$$R_n(\gamma) = E_{\boldsymbol{\lambda}}\left[L_n(\boldsymbol{\lambda}, \overline{\boldsymbol{Y}}_n)\right] = \frac{\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})}{n} + \gamma n. \quad (4)$$

Here, tr(.) denotes the trace of a matrix. By taking derivative of the expression in (4) one can find the n that will minimize the risk. That will give the optimal fixed-sample size as,

$$n^* \equiv n^*(\gamma) = [\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})/\gamma]^{1/2}, \quad (5)$$

which gives the minimum *fixed-sample risk* $R_{n^*}(\gamma) = 2\gamma n^*$. One should carefully note that, both $\boldsymbol{\lambda}$ and $\boldsymbol{\Sigma}$ are unknown, and therefore neither n^* nor $R_{n^*}(\gamma)$ can be determined. Hence, we adopt a purely sequential sampling procedure to estimate $\boldsymbol{P}\boldsymbol{\lambda}$. The sampling can be done according to the the stopping rule,

$$N \equiv N(\gamma) = \inf\left\{n \geq 2 : n \geq \sqrt{\frac{\text{tr}(\boldsymbol{W}\boldsymbol{S}_n)}{\gamma}}\right\} \quad (6)$$

where,

$$\boldsymbol{S}_n = \frac{1}{n-1}\sum_{k=1}^{n}(\boldsymbol{Y}^{(k)} - \overline{\boldsymbol{Y}}_n)(\boldsymbol{Y}^{(k)} - \overline{\boldsymbol{Y}}_n)^T$$

is the $r \times r$ sample dispersion matrix.

We now illustrate the algorithm to implement the stopping rule given in (6). First, we compute the mean vector $\bar{\boldsymbol{y}}_2$ and the sample-covariance matrix $\boldsymbol{s}_2 = \frac{1}{2-1}\sum_{k=1}^{2}(\boldsymbol{y}^{(k)} - \overline{\boldsymbol{y}}_2)(\boldsymbol{y}^{(k)} - \overline{\boldsymbol{y}}_2)^T$ from the two observed vectors of data $\boldsymbol{y}^{(1)}$ and $\boldsymbol{y}^{(2)}$. Note that $\boldsymbol{W}$ and γ are known to begin with. Now, we will compute the right-hand side of the inequality in (6) i.e. $[\text{tr}(\boldsymbol{W}\boldsymbol{s}_2)/\gamma]^{1/2}$ and check whether or not, $n = 2 \geq [\text{tr}(\boldsymbol{W}\boldsymbol{s}_2)/\gamma]^{1/2}$. If yes, then our estimate is $\bar{\boldsymbol{y}}_2$. If not, then we collect the next data vector, $\boldsymbol{y}^{(3)}$ and compute the updated mean vector $\bar{\boldsymbol{y}}_3$ and the co-variance matrix $\boldsymbol{s}_3$. One will again check from (6) whether or not, $n = 3 \geq [\text{tr}(\boldsymbol{W}\boldsymbol{s}_3)/\gamma]^{1/2}$. If yes, then our estimate will be $\bar{\boldsymbol{y}}_3$. If not, then we repeat the previous steps with the next data vector. This is how the algorithm will be continued until the condition to stop in (6) is satisfied. We will stop at the minimum value of n for which we observe, $n \geq [\text{tr}(\boldsymbol{W}\boldsymbol{s}_n)/\gamma]^{1/2}$ and hence take the observed value $\bar{\boldsymbol{y}}_n$ as our final estimator of the unknown vector $\boldsymbol{P}\boldsymbol{\lambda}$. Clerly, the stopping variable N is a random variable and we will explore some of its properties shortly.

Once the sampling procedure in (6) terminates and we estimate the parameter vector $\boldsymbol{P}\boldsymbol{\lambda}$ by the randomly

stopped mean vector $\overline{\boldsymbol{Y}}_N$. The associated sequential risk function is,

$$R_N(\gamma) = E_{\boldsymbol{\lambda}}\left[(\overline{\boldsymbol{Y}}_N - \boldsymbol{P}\boldsymbol{\lambda})^T\boldsymbol{W}(\overline{\boldsymbol{Y}}_N - \boldsymbol{P}\boldsymbol{\lambda})\right] + \gamma E_{\boldsymbol{\lambda}}\left[N\right]. \quad (7)$$

Since the estimation procedure is based on a stopping rule, it is important to verify that the procedure will stop eventually and also without much oversampling. The following theorem explains that.

Theorem 1. *For all fixed $\boldsymbol{\lambda}$ and $\boldsymbol{W}_{r \times r}$, for the the stopping rule in (6) we have the following properties,*

(i) $P_{\boldsymbol{\lambda}}[N < \infty] = 1$

(ii) $\lim_{\gamma \to 0} N n^{*-1} = 1$ *with probability 1.*

(iii) $\lim_{\gamma \to 0} E_{\boldsymbol{\lambda}}\left[N n^{*-1}\right] = 1$

where, n^* is defined in (5).

Proof. To prove *(i)*, first note that with probability 1, $\boldsymbol{S}_m \to \boldsymbol{\Sigma}$ as $m \to \infty$. Now,

$$\begin{aligned}P_{\boldsymbol{\lambda}}[N = \infty] &= \lim_{m \to \infty} P_{\boldsymbol{\lambda}}[N > m] \\ &\leq \lim_{m \to \infty} P_{\boldsymbol{\lambda}}\left[m < \left[\gamma^{-1}\text{tr}(\boldsymbol{W}\boldsymbol{S}_m)\right]^{1/2}\right] \\ &= 0.\end{aligned}$$

Next, from (5) and (6) we can claim, with probability 1,

$$\sqrt{\gamma^{-1}\text{tr}(\boldsymbol{W}\boldsymbol{S}_N)} \leq N < \sqrt{\gamma^{-1}\text{tr}(\boldsymbol{W}\boldsymbol{S}_{N-1})} + 1 \quad (8)$$

which implies

$$\sqrt{\frac{\text{tr}(\boldsymbol{W}\boldsymbol{S}_N)}{\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})}} \leq \frac{N}{n^*} < \sqrt{\frac{\text{tr}(\boldsymbol{W}\boldsymbol{S}_{N-1})}{\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})}} + \frac{1}{n^*}. \quad (9)$$

As $\gamma \to 0$, $N \to \infty$ and hence, $\boldsymbol{S}_{N-1} \to \boldsymbol{\Sigma}$ and $\boldsymbol{S}_N \to \boldsymbol{\Sigma}$ with probability 1. This proves *(ii)*.

For part (iii), we have from (8), with probability 1,

$$N n^{*-1} < \sqrt{\text{tr}(\boldsymbol{W}\boldsymbol{S}_{N-1})[\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})]^{-1}} + 1$$

$$< \sqrt{U[\text{tr}(\boldsymbol{W}\boldsymbol{\Sigma})]^{-1}} + 1 \quad (10)$$

where, $U = \max_{k \geq 1} \text{tr}(\boldsymbol{W}\boldsymbol{S}_{k-1})$. By ergodic theorem, all positive powers of U are integrable. Thus part (iii) follows from part (ii), (9) and dominated convergence theorem. $\square$

We also have a conjecture regarding the risk defined in (7). We first define the risk ratio as,

$$\eta(\gamma) = \frac{R_N(\gamma)}{R_{n^*}(\gamma)} = \frac{R_N(\gamma)}{2\gamma n^*}. \quad (11)$$

Conjecture 1. *For all fixed $\boldsymbol{\lambda}$ and $\boldsymbol{W}_{r \times r}$, we have,*

$$\lim_{\gamma \to 0} \eta(\gamma) = 1 \tag{12}$$

where, n^* is defined in (5) and $\eta(\gamma)$ is defined on (11).

Theorem 1 along with Conjecture 1 essentially claims that the estimation strategy is efficient in estimating the parameter vector $\boldsymbol{P\lambda}$. Equation (12) is actually claiming the *risk-efficiency* of the procedure. While we postpone the rigorous proof of this claim, an empirical verification of (12) will be provided in Section 4.

4 Data Analysis

In this section we discuss the performance of our estimation procedure discussed in Section 3. We assume that $\boldsymbol{P}$ is known to us either directly or through the random-routing matrix $\boldsymbol{A}$. For brevity, let us consider $\boldsymbol{W} = \boldsymbol{I}_{r \times r}$. Hence, from (5) we have,

$$n^* = [\mathrm{tr}(\boldsymbol{\Sigma})/\gamma]^{1/2} = \gamma^{-1/2} \sum_i \sum_j \lambda_j P_i^j. \tag{13}$$

Here we revisit the example discussed in [1]. A simple network is described in Figure 1 where we have 4 communicating nodes, a, b, c and d, forming $4 \times 3 = 12$ SD pairs. In addition, these nodes communicate via 7 directed links. These nodes communicate via specific links which can be represented by the *random Markovian-routing* matrix $\boldsymbol{A}$ given below. Cell with 0 as the entry is kept blank.

	1 ab	2 ac	3 ad	4 ba	5 bc	6 bd	7 ca	8 cb	9 cd	10 da	11 db	12 dc
1($a \to b$)	0.8	0.2	0.2									
2($a \to c$)	0.2	0.8	0.8		1.0	1.0						
3($b \to a$)				1.0	0.2	0.1	1.0			1.0		
4($b \to c$)		0.8			0.8	0.1						1.0
5($b \to d$)		0.2	1.0			0.8			1.0			
6($c \to b$)	0.8		0.2				0.8	0.8	0.2	1.0	1.0	
7($c \to d$)	0.2		0.8			1.0	0.2	0.2	0.8			
8($d \to b$)	1.0						1.0	1.0		0.8	0.8	0.2
9($d \to c$)		1.0								0.2	0.2	0.8

(matrix $\boldsymbol{A}$)

The entries of $\boldsymbol{A}$ are conditional probabilities, already defined in Section 2. As an example, we consider a specific SD pair ad and the links involved in the communication are shown in Figure 1.
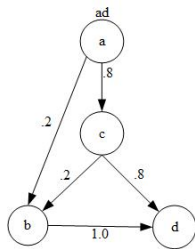

Figure 1: Four node directed graph.

Now suppose we want to calculate $P_3^7 (= P_{c \to d}^{ad})$, the probability that a message with SD address ad passes through link $c \to d$. The general multiplication rule of probabilities says that $P(ABC) = P(A|BC)P(B|C)P(C)$. Noting that the entries of $\boldsymbol{A}$ are conditional probabilities, applying the general multiplication rule we get, $P_3^7 = a_{23} \times a_{73} = 0.8 \times 0.8 = 0.64$. Similar calculations give the other entries of the matrix $\boldsymbol{P}$ as shown below. As in matrix $\boldsymbol{A}$, a cell with 0 as the entry is kept blank.

	1 ab	2 ac	3 ad	4 ba	5 bc	6 bd	7 ca	8 cb	9 cd	10 da	11 db	12 dc
1($a \to b$)	0.8	0.2	0.2									
2($a \to c$)	0.2	0.8	0.8		0.2	0.1						
3($b \to a$)				1.0	0.2	0.1	1.0			1.0		
4($b \to c$)		0.16			0.8	0.1						0.2
5($b \to d$)		0.04	0.36			0.8			0.2			
6($c \to b$)	0.16		0.16				0.8	0.8	0.2	0.2	0.2	
7($c \to d$)	0.04		0.64			0.2	0.2	0.2	0.8			
8($d \to b$)	0.04						0.2	0.2		0.8	0.2	0.2
9($d \to c$)		0.04								0.2	0.2	0.8

(matrix $\boldsymbol{P}$)

Here we consider an arbitrary vector

$$\boldsymbol{\lambda} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)^T$$

to be the true parameter and hence with known matrix $\boldsymbol{P}$, we have

$$\boldsymbol{P\lambda} = (1.8, 5.8, 22.6, 7.32, 15.68,$$
$$18.64, 17.32, 19.6, 21.8)^T.$$

To perform the simulations, we start by generating random observations of the vector $\boldsymbol{Y} = (Y_1, Y_2, ..., Y_9)^T$ one by one from the Poisson distribution with mean vector $\boldsymbol{P\lambda}$. With each new sample of vector, we check the stopping rule (6). This process was also described in the paragraph following equation (6). Suppose the first time, the randomly stopped sample size was $N_1 = n_1$. Hence we have the first estimate of the true $\boldsymbol{P\lambda}$ as $\widehat{\boldsymbol{P\lambda}}_1 = n_1^{-1} \sum_{i=1}^{n_1} \overline{\boldsymbol{Y}}_{n_1}$ for $\boldsymbol{P\lambda}_1 = \sum_{j=1}^c \lambda_j P_1^j$. This will be the first replication of the process. If this entire process was replicated R times, we have values, N_i, $\widehat{\boldsymbol{P\lambda}}_i$ and $R_{N_i}(\gamma)$ for $i = 1, 2, ..., R$. Table 1 and Table 2 present the summary of these statistics for $R = 10000$ simulations. Let us first define the following quantities:

$$\overline{N} = R^{-1} \sum_{i=1}^R N_i, \quad \widehat{\boldsymbol{P\lambda}} = R^{-1} \sum_{i=1}^R \widehat{\boldsymbol{P\lambda}}_i,$$
$$\overline{R}_N(\gamma) = R^{-1} \sum_{i=1}^R R_{N_i}(\gamma) \text{ and}$$
$$\hat{\eta}(\gamma) = \overline{R}_N(\gamma)/R_{n^*}(\gamma).$$

Column 1 of Table 1 shows different values of the cost factor γ. As γ decreases, the value of N grows as expected. The second and the third columns of Table 1 are revealing the behavior of the stopped sample size. It is quite encouraging to see that the ratio $\overline{N}/n^*$ remains close to 1 even for smaller values of $\overline{N}$. The

Table 1: Results of Simulation

γ	$\overline{N}$	$\overline{N}/n^*$	$\overline{R}_N(\gamma)$	$\hat{\eta}(\gamma)$
0.1	33.1114	1.0132	6.5429	1.0010
0.05	46.6402	1.0092	4.6259	1.0009
0.01	103.7709	1.0041	2.0714	1.0022
0.005	146.6068	1.0031	1.4630	1.0010
0.001	327.2686	1.0014	0.6562	1.0039

Table 2: Estimated values of $P\lambda$ for given γ

$P\lambda$	values of γ				
	0.1	0.05	0.01	0.005	0.001
1.80	1.7968	1.7976	1.7970	1.7975	1.7998
5.80	5.7968	5.8003	5.7997	5.7979	5.7992
22.60	22.5918	22.5918	22.5978	22.5985	22.6046
7.32	7.3203	7.3211	7.3206	7.3231	7.3226
7.76	7.7659	7.7596	7.7627	7.7605	7.7581
18.64	18.6399	18.6401	18.6353	18.6408	18.6401
13.36	13.3576	13.3547	13.3650	13.3620	13.3576
15.64	15.6369	15.6331	15.6413	15.6398	15.6411
13.88	13.8717	13.8736	13.8781	13.8804	13.8801

fourth and the fifth columns correspond to the values of simulated risk and the ratios of simulated risk over optimal risk. The efficiency of the estimation procedure is demonstrated by the values of $\hat{\eta}(\gamma)$ which are close to 1 even for smaller values of N.

Table 2 shows the estimated values of the true $P\lambda$ for the different choices of γ from Table 1. The true value and estimated values are extremely close for every single choice of γ. Estimates for other choices of λ were equally good. We do not present them for brevity.

5 Conclusion and Future Directions

In this paper, we proposed an innovative and efficient method of estimating a vector of parameters that characterize the node-to-node traffic intensities in a network. The error in the method of estimation is clearly defined by a risk function. A stopping rule is developed which determines the required number of observations to provide an estimate of the parameters while minimizing risk (error) in estimation. Unlike ML/EM procedures, our process is unaffected by any good or bad choice of initial estimators. Moreover, it saves the cost of estimation in terms of money, time or other logistics, without oversampling on average. The efficiency of the adaptive estimation process is demonstrated by theoretical derivations and validated through extensive simulations. The estimation of $P\lambda$ is highly accurate even when the sample size is small.

One may note that the estimation is conducted under a Poisson probability model. In future research, we will work on estimating $P\lambda$ under different distribution models or even with a distribution-free setup.

References:

[1] Vardi, Y. (1996). Network tomography: Estimating source-destination traffic intensities from link data, *Journal of the American statistical association*, 91: 365-377.

[2] Khan, MSS and Kumar, Anup and Xie, Bin and Sahoo, Prasanna K. (2015). Network tomography application in mobile ad-hoc network using stitching algorithm, *Journal of Network and Computer Applications*, 56: 77-87.

[3] Paranjothi, Anirudh and Khan, Mohammad S and Patan, Rizwan and Parizi, Reza M and Atiquzzaman, Mohammed. (2020). VANETomo: A congestion identification and control scheme in connected vehicles using network tomography, *Computer Communications*.

[4] Chiu, Cho-Chun and He, Ting. (2020). VStealthy DGoS Attack: DeGrading of Service under the Watch of Network Tomography, *IEEE INFOCOM*.

[5] Tsang, Yolanda and Coates, Mark and Nowak, Robert D. (2003). Network delay tomography, *IEEE Transactions on Signal Processing*, 51: 2125-2136.

[6] Sharma, G. and Jaggi, S. and Dey, B. K. (2008). Network tomography via network coding, *2008 Information Theory and Applications Workshop*, pp. 151-157.

[7] Kavacky, Matej. (2018). Software IMS Core Network under SIP Load Evaluation, *International Journal of Circuits, Systems and Signal Processing*, 12: 684-689.

[8] Itradat, Awni and Mnayyes, Abdullah. (2019). A scheme for identifying and correcting the network weaknesses by incorporating bandwidth measurements, *International Journal of Circuits, Systems and Signal Processing*, 13: 520-529.

[9] Snyder, D. L., Schultz, T. J., and O'Sullivan, J. A. (1992). Deblurring subject to nonnegativity constraints, *IEEE Transactions on Signal Processes*, 40: 1143-1150.

[10] Vardi, Y., and Lee, D. (1993). From Image Deblurring to Optimal Investments: Maximum likelihood solutions for positive linear inverse problems (with discussion), *Journal of the Royal Statistical Society, Series B*, 55: 569-612.

[11] Biersack, Ernst and Callegari, Christian and Matijasevic, Maja (2013). *Data traffic monitoring and analysis: from measurement, classifica-*

*tion, and anomaly detection to quality of experi-
ence*, Springer.

[12] Mukhopadhyay, N. and de Silva, B. M. (2009).
Sequential methods and their applications, CRC.

[13] Rovatsos G., Zou S. and Veeravalli V. V. (2020).
Sequential algorithms for moving anomaly detec-
tion in networks, *Sequential Analysis*, 39: 6-31.

[14] Ghosh, M. and Mukhopadhyay, N. (1981). Con-
sistency and asymptotic efficiency of two-stage
and sequential procedures, *Sankhyā, Series A*, 43:
220-227.

[15] Hall, P. (1983). Sequential estimation saving
sampling operations, *Journal of Royal Statistical
Society, Series B*, 45: 219-223.

[16] Mukhopadhyay, N. and Zhuang, Y. (2019),
Two-sample two-stage and purely sequential
methodologies for tests of hypotheses with appli-
cations: comparing normal means when the two
variances are unknown and unequal, *Sequential
Analysis*, 38: 70-115.

[17] Mukhopadhyay N. and Bhattacharjee D. (2018).
Sequentially estimating the required optimal ob-
served number of tagged items with bounded risk
in the recapture phase under inverse binomial
sampling, *Sequential Analysis*, 37: 412-429.

[18] Zhuang Y. and Bapat S. (2020). On Compar-
ing locations of two-parameter exponential dis-
tributions using sequential sampling with appli-
cations in cancer research, *Communication in
Statistics- Simulation and Computation*. DOI:
10.1080/03610918.2020.1794007

[19] Hu, J. and Mukhopadhyay, N. (2019). Second-
order asymptotics in a class of purely sequential
minimum risk point estimation (MRPE) method-
ologies. *Japanese Journal of Statistics and Data
Science*, 2: 81-104.

[20] Zhuang Y. and Bhattacharjee D. (2020). Min-
imum risk point estimation of the size of a fi-
nite population under mark-recapture strategy,
Sequential Analysis. In press.

Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

Debanjan Bhattacharjee developed the main method
of estimation. Yan Zhuang carried out the simula-
tions. Bhattacharjee and Zhuang work together in
writing the paper.