Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Technology 3 (2012) 41 – 51
Procedia Technology

Stitching Algorithm: A Network Performance Analysis Tool for Dynamic Mobile Networks

M. S. S. Khan[a,*], Anup Kumar[a], Bin Xie[b]

[a]Duthie Center of Engineering
Department of Computer Engineering and Computer Science
University of Louisville, Louisville, KY 40292, USA.
[b]InfoBeyonds Technology LLC, Louisville, KY, USA

Abstract

The performance analysis of the mobile ad-hoc network (MANET) is a challenging issue. In this paper, network tomography is studied to analyse the network performance in a dynamic MANET. For such a purpose, a network tomography analytical model is proposed for a dynamic network environment. Expected Maximization (EM) algorithm for network tomography is able to estimate the network performance parameter in accordance to network performance observations. Our study is different than current network tomography approaches where are applied for static wired network. Over the dynamic network, we proposed a new algorithm that is called Stitching algorithm to aggregate the dynamic performance. Specifically, the stitching algorithm concatenates the performance parameter i.e. link delay, from distinguish time periods. Therefore, the network behaviour as well as the corresponding performance in a mobile ad-hoc network can be derived over a continuous period.

Keywords: Mobile Networks, Expected Maximization, Maximum Likelihood Estimator, Stitching Algorithm.

1. Introduction

Recently, the network tomography (NT) is proposed for studying the network performance. One of the initial works is conducted by Vardi [1] in 1996 and several algorithms [2] [6] are developed for analyzing the network performance. These proposed network tomography algorithms are based on the Internet or static computer networks. Therefore, these network tomography algorithms are not directly applicable for dynamic network such as mobile wireless ad hoc networks (MANET). The estimation accuracy of existing network tomography models is not acceptable for Mobile Ad-hoc network since these models fail to addresses the mobile Ad-hoc network topology. In other words, these models have to be improved with the consideration of high mobility. In this paper, we investigate the network tomography to study the dynamic network performance. A mathematical model and Pseudo-log Likelihood Estimation (PLE) algorithm are developed to offer a multilevel monitoring and evaluation of network performance parameters.

In our work, PLE is a network tomography approach to reveal the probabilistic distribution of network performance parameters, such as route delay, e.g., how the end-to-end packet delay varies in the network

*Corresponding author
Email address: mskhan02@louisville.edu (M. S. S. Khan)

2212-0173 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.protcy.2012.03.005

with time. PLE performs network tomography analysis in a Mobile ad-hoc network. This is different than the current approaches, which estimate the network performance parameters on the Internet or static computer networks. In a Ad-hoc mobile network PLE is able to monitor, analyse, and manage the information flows in such a way the network behaviour can be tracked in different time scales. In other words, the network performance can be evaluated and predicted for given network conditions.

As the increase of the ad-hoc mobile network mobility, this period of steady state generally decreases. For each steady state, however, PLE is able to perform the estimation of network performance parameters. Along with the transition of Ad-hoc mobile network, PLE then aggregates the results from multiple continuous steady states. The aggregated results indicate the network performance variations over the time domain. The organization of the paper is as follows: Section 2 discusses the mathematical model for NT problem. Section 3 talks about EM algorithm and some simulation results. In section 4 we talk about the stitching algorithm. Results, conclusion and acknowledgement are given in sections 5, 6 and 7.

2. NT Mathematical Model

Let us consider the following eight node network. Matrix A represents the path matrix of the network in Fig 1.

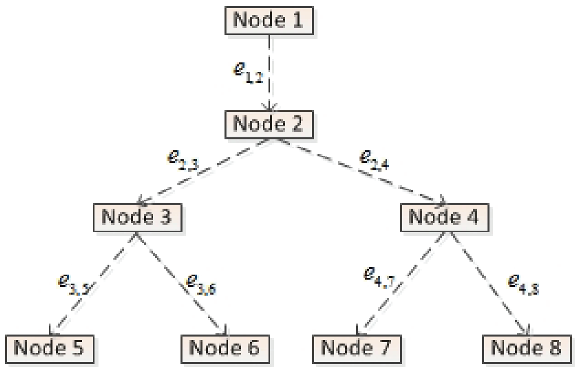


Fig. 1: Eight-Node Directed Network in a Given Time Instant

In matrix A, 1 means the sink is connected with the node on the topology while 0 means otherwise. In matrix A, 1 means the link is connected with the node on the topology while 0 means otherwise:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Here $\mathbf{X} = (x_1, x_2, ..., x_m)^t$ be the m dimensional network performance random vector that present a network performance parameter such as link delay, packet delay or link available bandwidth. The superscript t is the transpose operation. Thus $x_{i,1<i<m}$ is the quantitative evaluation of the i^{th} link performance for the information flow. Let $\mathbf{Y} = (x_1, x_2, ..., x_n)^t$ be the n^{th} dimensional measurement vector corresponding to $\mathbf{X}$ and $y_{i,1<i<n}$ is the observed i^{th} link performance for the information flow. The problem of the network tomography is to estimate the distribution probability of $\mathbf{X}$ from the observed $\mathbf{Y}$ which can be modelled as a matrix equation :

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_7 \end{pmatrix} \tag{1}$$

Given x_j that is the j^{th}component of $\mathbf{X}$, we assume our data follows Poisson distribution. Then the probability mass function(pmf)is given by

$$p(x;\theta) = \frac{\theta^x e^{-\theta}}{x!}, x = 0, 1, 2, \ldots \tag{2}$$

Here $\theta \in \mathbb{R}^+$,which is equal to the expected number of occurrences that happen during the given interval and x is the number of occurrences of an event, the probability of which is given by (2). Considering all the components in $\mathbf{X}$, it has $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\}$ that is the parameter of the whole model. The observed data vector $\mathbf{Y}$ for subsequently observed vectors for $\boldsymbol{T}$ intervals and denote $\boldsymbol{X}_t, t = 1, 2, \ldots, \boldsymbol{T}$ as the unknown network performance quantities corresponding to $\mathbf{Y}_t$. Furthermore, $\mathbf{Y}_{t,i}$ and $\mathbf{X}_{t,j}$ be the i^{th} and j^{th} element of $\mathbf{Y}_t$ and $\mathbf{X}_t$ respectively.We first consider $\mathbf{Y}_t$ is collected from a stable network topology graph (e.g., the multicast topology for a certain period). For example, a multicast topology in the network may maintain its multicast topology for a certain period and in this period we have the corresponding observed $\mathbf{Y}$ in this period is $\boldsymbol{Y}_t, t = 1, 2, \ldots, \boldsymbol{T}$. For this purpose, the raw data from the dynamic network should be preprocessed according to the topology. It is more convenient to work in terms of the natural logarithm of the likelihood function, called the log-likelihood, than in terms of the likelihood function itself. Because the logarithm is a monotonically increasing function, the logarithm of a function achieves its maximum value at the same points as the function itself.

A Pseudo-likelihood is an approximation to the joint probability distribution of a collection of random variables. The practical use of this is that it can provide an approximation to the likelihood function of a set of observed data which may either provide a computationally simpler problem for estimation, or may provide a way of obtaining explicit estimates of model parameters. So in (3), we have the sum of all the random variables $\mathbf{Y}$ over the parameter θ. Also pseudo-likelihood in place of the true likelihood function in a maximum likelihood analysis can lead to good estimates. In practice we may have more than one $\mathbf{Y}$. So to approximate θ over all the $\mathbf{Y}$ will yield better results than approximating θ for each $\mathbf{Y}$. This may not be an issue, when dealing with few $\mathbf{Y}$ but in problems, where we have many $\mathbf{Y}$. The pseudo log-likelihood function will give better result.We now define our pseudo-log likelihood functions:

$$L_T^p(Y_1, Y_2, \ldots, Y_T; \theta) = \sum_{t=1}^{T} L^p(Y_t; \theta) \tag{3}$$

$$L^p(Y_t; \theta) = \log p(Y_t; \theta) \tag{4}$$

here $p(Y_t; \theta)$ is the marginal likelihood function. To maximize the likelihood estimate of parameter θ is to maximize the likelihood function in(3);

$$\frac{\partial}{\partial\theta} L_T^p(T_1, T_2, \ldots, Y_T; \theta) = 0 \tag{5}$$

The likelihood equation in (5) has a unique solution almost surely as $T \longrightarrow \infty$. Therefore, it needs to select big enough samples from the dynamic network data to have a good estimate. Since we are dealing with a big dataset, this approach is not appropriate to solve (5) directly. That is because we have defined (5) in terms of marginalized likelihood function, which is difficult to compute directly. Instead, we need to adopt a numerical optimization technique, which will solve for $\mathbf{Y}$ iteratively. This give rise to a very well know algorithm called Expectation-Maximization or more commonly known as EM algorithm.

3. Expected Maximization Algorithm

Given a statistical model consisting of a set $\mathbf{Y} = A\mathbf{X}$ observed data, a set of unobserved data $\mathbf{X}$ a vector of unknown parameters θ , along with a likelihood function $L^p(Y_t; \theta) = \log p(Y_t; \theta))$, the maximum likelihood estimate (MLE) of the unknown parameters is determined by the marginal likelihood of the observed data

$$L^p(Y_t; \theta) = \log p(Y_t; \theta) = \sum_{t=1}^{T} L^p(Y_t; \theta) \tag{6}$$

In (6), the right hand summation represents the marginal likelihood function, where θ is the parameter of interest with respect to $\mathbf{Y}$ and $\mathbf{X}$. However, θ is intractable. Therefore, we define the objective function $Q\left(\theta, \theta^i\right)$ to be maximized over θ in the $(i+1)^{th}$ step of the EM algorithm. The EM algorithm seeks to find the MLE of the marginal likelihood function iteratively by applying the following two step:

Expectation Step (E-step): Calculate the expected value of the log likelihood function, with respect to the conditional distribution of $\mathbf{X}$ given $\mathbf{Y}$ under the current estimate of the parameters $\theta^{\mathbf{i}}$:

$$Q\left(\theta, \theta^i\right) = E_{X|Y,\theta^i} \log p\left(Y_t; \theta\right)$$

Maximization Step (M-step): Find the parameter that maximizes this quantity:

$$\theta^i = \underset{\theta}{argmax} Q\left(\theta, \theta^i\right) \tag{7}$$

The pseudo code of the algorithm is given below.

Algorithm 1 Expected Maximization Algorithm

begin
initialize $\theta^0, \epsilon, i \leftarrow 0$
do $i \leftarrow i+1$
E step: compute $Q(\theta, \theta^i)$
M step: $\theta^{i+1} \leftarrow \underset{\theta}{argmax} Q\left(\theta, \theta^i\right)$
until $Q(\theta^{i+1}; \theta^i) - Q(\theta^i; \theta^{i-1}) \leqslant \epsilon$
return $\hat{\theta} \leftarrow \theta^{i+1}$
end

3.1. Validation of EM Algorithm

We now present the details , as to how the EM algorithm works in the network tomography problem. Let us consider the following topology.

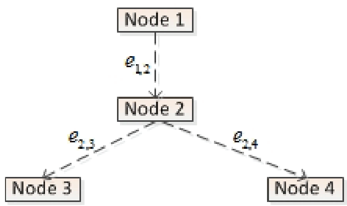


Fig. 2: Four-Node Directed Network

Then we use the matrix equation $\mathbf{Y} = A\mathbf{X}$ to represent the Fig 3, this is shown in (8). We assume, $X_i \sim Poission\left(\theta_i\right)$.

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \tag{8}$$

In (8), matrix A represents the deterministic connectivity between the two nodes in a given topology. So, from (8), we have the following two equations:

$$1 = X_1 + X_2 \tag{9}$$

$$2 = X_1 + X_2 \tag{10}$$

So (9) has two solutions $(1, 0, 1)$ and$(0, 1, 2)$.We claim for any arbitrary given θ our scheme will converge to the MLE. We drive the likelihood equations to prove our claim. Let

$$L^p(Y;\theta) = P_\theta\left\{Y = (1,2)^{'}\right\} \tag{11}$$

$$= P_\theta\left\{X = (1,0,1)^{'}\right\} + P_\theta\left\{X = (0,1,2)^{'}\right\} \tag{12}$$

$$= \left(\theta_1\theta_3 + \frac{\theta_2\theta_3^2}{2!}\right)\exp(-\theta_1 - \theta_2 - \theta_3) \tag{13}$$

$$= \left(\frac{2\theta_1\theta_3 + \theta_2\theta_3^2}{2}\right)\exp(-\theta_1 - \theta_2 - \theta_3) \tag{14}$$

Since we are computing the probability of **Y** based on Poisson distribution, from(14) we can write the following

$$\frac{\theta_3}{2}(2\theta_1 + \theta_2\theta_3)\exp(-\theta_1 - \theta_2 - \theta_3) = 1 \tag{15}$$

$$\Rightarrow \exp(-\theta_1 - \theta_2 - \theta_3) = \frac{2}{\theta_3(2\theta_1 + \theta_2\theta_3)} \tag{16}$$

Therefore, probability of each **X** is given by

$$p(X = (1,0,1)^{'}) = \theta_1\theta_3\exp(-\theta_1 - \theta_2 - \theta_3) \tag{17}$$

$$= \theta_1\theta_3\left(\frac{2}{\theta_3(2\theta_1 + \theta_2\theta_3)}\right) \tag{18}$$

$$= \frac{2\theta_1}{2\theta_1 + \theta_2\theta_3} \tag{19}$$

Similarly, we can compute

$$p(X = (0,1,2)^{'}) = \theta_1\theta_3\exp(-\theta_1 - \theta_2 - \theta_3) \tag{20}$$

$$= \frac{2\theta_2\theta_3}{2\theta_1 + \theta_2\theta_3} \tag{21}$$

Hence, by first principle

$$E[X|Y,\theta] = \sum_{i=1}^{2} x_i p(x_i) \tag{22}$$

$$= \begin{pmatrix}1\\0\\1\end{pmatrix}\frac{2\theta_1}{2\theta_1 + \theta_2\theta_3}\begin{pmatrix}0\\1\\2\end{pmatrix}\frac{2\theta_2\theta_3}{2\theta_1 + \theta_2\theta_3} \tag{23}$$

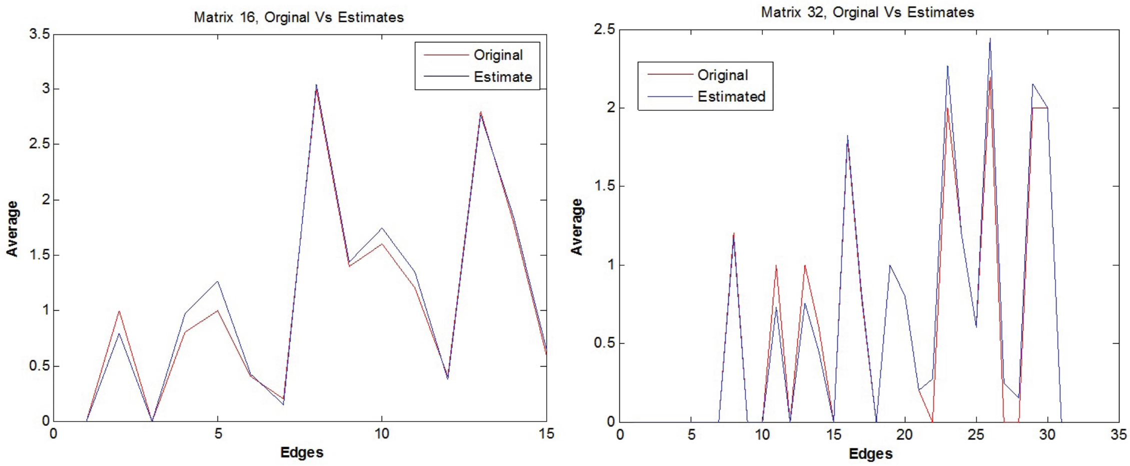
We can write (23) in the following fashion

$$\begin{pmatrix}\theta_1\\\theta_2\\\theta_3\end{pmatrix} \leftarrow E[X|Y,\theta] = \begin{pmatrix}\frac{2\theta_1}{2\theta_1+\theta_2\theta_3}\\\frac{2\theta_2\theta_3}{2\theta_1+\theta_2\theta_3}\\\frac{1+\theta_2\theta_3}{2\theta_1+\theta_2\theta_3}\end{pmatrix} \tag{24}$$

For any arbitrary choice of $\theta = (\theta_1, \theta_2, \theta_3)$, we plug it back in (22), till the difference of two consecutive iteration is minimum,i.e. we find no chance in the iteration scheme output. So, we get the following result.

$$\begin{pmatrix}0.6\\0.3\\1.3\end{pmatrix} \Rightarrow \begin{pmatrix}0.75\\0.24\\1.24\end{pmatrix} \Rightarrow \ldots \begin{pmatrix}0.99\\0.01\\1.01\end{pmatrix} \Rightarrow \begin{pmatrix}1\\0\\1\end{pmatrix} \Rightarrow \begin{pmatrix}1\\0\\1\end{pmatrix} \tag{25}$$

In a similar fashion we computed the MLE based on 16 and 36 node network.



(a) 16 node network EM estimation result

(b) 32 node network EM estimation result

Fig. 3: EM Algorithm Original and Estimated edge values for various Network sizes

It can be clearly seen in the above figures that our estimates based on the EM algorithm are good.

4. Stitching Alorithm

The proposed stitching algorithm is a network performance evaluation algorithm for network topology that is changing in time. The basic idea of the stitching algorithm is to observe some characteristic of the topology from which one is able to identify the time variation. As most of the dynamical system, one underlying factor of a dynamic network is the stability analysis. Thus, knowing the behavior of the dynamic network in terms of stability can be used to determine various time intervals. In our study, we have the observed topology graph, which governs the topology of the network in time. Thus, knowing the underlying behavior or characteristic can help us to determine various time slots. One important aspect of these network graphs is they can be very easily translated into a path matrix. The stitching algorithm understands the performance variation corresponding to the topology change.

We now define an isomorphism $\Phi : H \to V$, where H and V are given matrices of network topologies. Then, H and V are same if only if following holds:

$$\Phi(H)_{size} = \Phi(V)_{size} \tag{26}$$

$$\Phi(H - V) = 0_{matrix} \tag{27}$$

If the conditions (26) and (27) are not true, we know the matrix has changed and hence the network topology has changed as well. When the network topology changes between measurement instances t_4 and t_5, we cannot take a simple average across observation periods. In this case, the stitching of two topologies has to take place. The stitching is performed in such a manner that the performance parameters of common edge values are averaged and others are kept as it is. The stitching procedure is to average the common edges of the network topologies while the remaining edges are concatenated into a new vector, representing as new *T's*. To understand the way *T's* work, we illustrate the procedure in Fig 4.

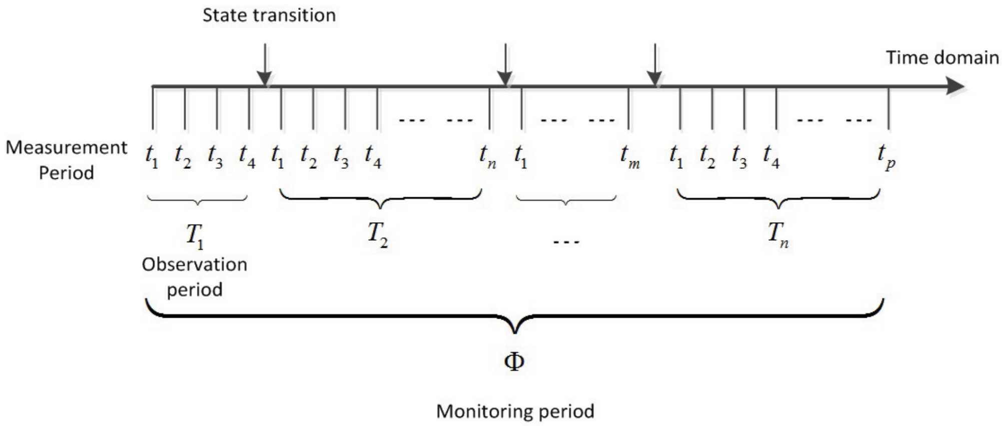


Fig. 4: Time Domain Parameter Measurement of Dynamic Network

Suppose the topology graph in Fig 5(a) represents connectivity at measurement time T_1 while topology graph in Fig 5(b) represents connectivity at measurement time T_2. As we can observe from the two graphs, edges $e_{1,2}, e_{2,3}, e_{2,4}$ are common to both topologies and others are not. In calculating combined network performance parameters for both graphs, the average of $e_{1,2}, e_{2,3}, e_{2,4}$ will be taken while parameters values of other edges will be added to the edge set. The edge set for observation period T_1 can be expressed as:

$$Edge\ set(T_1) = [e_{1,2}, e_{2,3}, e_{2,4}] \tag{28}$$

The edge set for the observation period T_2 can be expressed as:

$$Edge\ set(T_2) = [e_{1,2}, e_{2,3}, e_{2,4}, e_{3,5}, e_{3,6}, e_{4,7}, e_{4,8}] \tag{29}$$

Then the new T_2 is represented by the following stitching operation:

$$\Phi = [Edge\ set(T_1) \cup Edge\ set(T_2)] \tag{30}$$
$$= [e_{1,2}(T_1 \cup T_2), e_{2,3}(T_1 \cup T_2), e_{2,4}(T_1 \cup T_2), e_{3,5}, e_{3,6}, e_{4,7}, e_{4,8}] \tag{31}$$

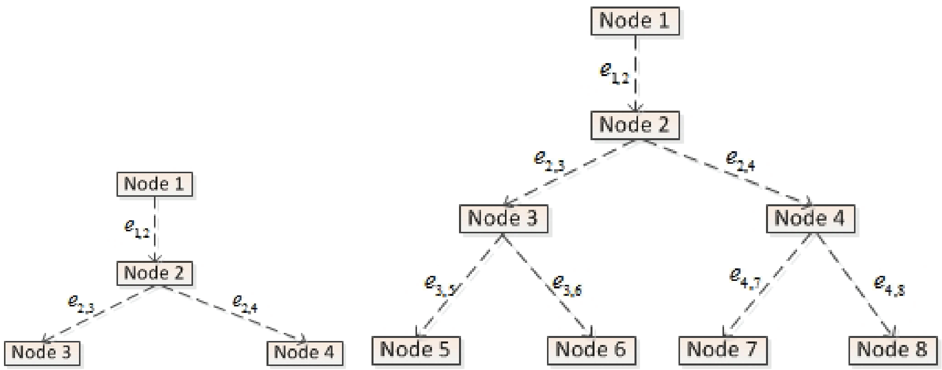
here $e_{i,j}(T_1 \cup T_2)$ is the stitching operation for two observation periods. We define $e_{i,j}(T_1 \cup T_2)$ as

$$e_{i,j}(T_1 \cup T_2) = \frac{\sum_{i=1}^{n_1} e_{i,j}(T_1, t_i) + \sum_{i=1}^{n_2} e_{i,j}(T_2, t_i)}{(n_1 + n_2)} \tag{32}$$

Here $e_{i,j}(T_1, t_i)$ is the performance evaluation for t_i during observation period T_1 and similarly $e_{i,j}(T_2, t_i)$ is the performance evaluation for t_i during the T_2. Also n_1 and n_2 are the number of measurement periods for T_1 and T_2 respectively. In (32), we assume that $i \neq j$. As shown in (31), only the common link performance parameters are naturally aggregated with the previous evaluations.

4.1. Simulation Results

We now present two cases, which represent variations of applying stitching algorithm. Stitching is applied on network presented in Fig 5(a) and 5(b).



(a) 4 node network at T_1 (b) 8 node network at T_2

Fig. 5: Directed Networks in a Given Time Instant

Case 1: In this case we consider the topology is unchanged from one time instance to another. Assume that the current state of topology is of size 4 nodes and next topology is also of size 4 from Fig 5(a).

Table 1: Estimates for 4 node graph

Observations	Edges		
	$e_{1,2}$	$e_{2,3}$	$e_{2,4}$
Original	2	2	0
Estimates	1.9980	2.0020	0.0020

Table 2: Estimates for 4 node graph

Observations	Edges		
	$e_{1,2}$	$e_{2,3}$	$e_{2,4}$
Original	7	0	13
Estimates	6.9860	0.0140	13.0140

We again enter another 4 node graph and its observations are presented in Table 2.Now, in the third time instance we expect the topology to change. So as expected, the above to results are aggregated and presented in the table below. Now, in the third time instance we expect the topology to change. So as expected, the above to results are aggregated and presented in the table below.

Table 3: Average Estimates

Observations	Edges		
	$e_{1,2}$	$e_{2,3}$	$e_{2,4}$
Average Original	4.5	1	6.5
Average Estimates	4.492	1.008	6.508

Case 2: Now suppose, next topology is changed and it is a 8 node graph as presented in Fig 5(b). So according to our algorithm, values in table 3 will be added with the common edges and then aggregated. The remaining edge values are concatenated as show in table 4.

In Fig 6, we compare Estimates using Stitching algorithm with the Original values.

Table 4: New average estimates with the concatenated values.

Observations	Edges						
	$e_{1,2}$	$e_{2,3}$	$e_{2,4}$	$e_{3,5}$	$e_{3,6}$	$e_{4,7}$	$e_{4,8}$
Average Original	4.5	1	6.5				
Original	4	0	4	2	0	0	4
Average Estimates	4.492	1.008	6.508				
Estimates	3.8652	0.1339	4.0078	2.0009	0.0009	0.1269	4.1269
New Average Estimates	4.1786	0.5709	5.2579	2.0009	0.0009	0.1269	4.1269
New Average Original	4.25	0.5	5.25	2	0	0	4

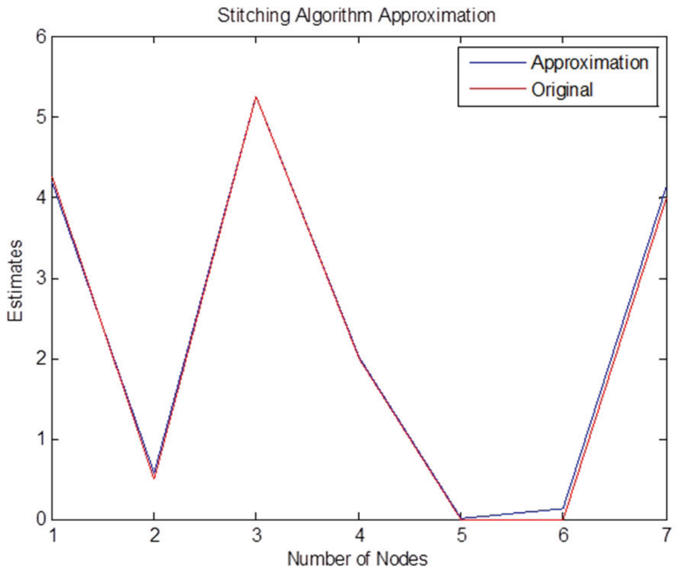


Fig. 6: Comparison between the Estimated and the Actual values using Stitching algorithm

5. Conclusion

Poisson based EM algorithm was implemented to find the MLE of an expected log likelihood function. The simple nature of the algorithm is a good tool to understand network estimation problems. This can be seen very clear from our numerical results. We assumed the Poisson distribution since the information is transmitted in a time instances albeit the time between the two transmissions could very small or reasonably long. To the center of this paper is the proposed Stitching algorithm. It is a very simple but yet a very elegant way to monitor network performance over the changing topology. It was evident from our simulation result that Stitching algorithm produce good results.

6. Acknowledgement

This research is supported by US Air Force grant AF9550-10-C-0153.

Appendix A.

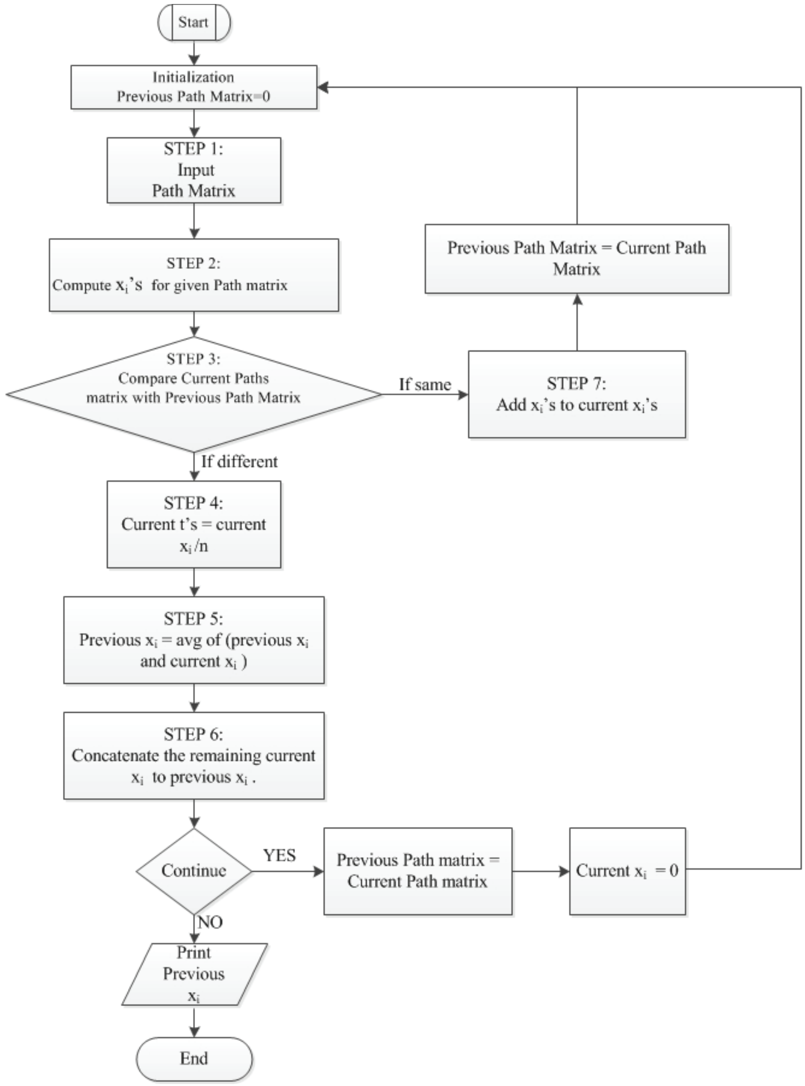


Fig. A.7: Flow chart of the Stitching Algorithm

References

[1] Y. Vardi. Network Tomography: Estimating source-destination traffic intensities from link data, *Journal of the American Statistical Association*,1996 ,91(433), 365-377.
[2] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu. Internet Tomography: Recent Developments, *Statistical Science*, 2004, 19(3), 499-517.
[3] G. Liang and B. Yu. Maximum Pseudo Likelihood Estimation in Network Tomography, *IEEE Trans. on Signal Processing (Special Issue on Data Networks)*, 2003, 51(8), 2043-2053.
[4] R. Caceres, N.G.Duffield, J. Horowitz and D. Towsley. Multicast Based Inference of Network Internal Loss Characteristics, *IEEE Trans. Information Theory*,1999 ,45 , 2462-2480.
[5] F. Lo Presti, N. G. Duffield, J. Horowitz, D. Towsley. Multicast-based Inference of Network-internal Delay Distributions, *IEEE/ACM Trans on Networking*,2002, 10,761-775.
[6] Ye Yao and Wandong Cai. Ad Hoc Network Measurement Based on Network Tomography: Theory, Technique and Application, *Journal of Networks*,2010 ,5(6).
[7] Y.Tsang, M. Coates, R.D. Nowak. Network Delay Tomography, *IEEE Trans. on Signal Processing, special issue on Signal Processing in Networking Data Networks)*,2003 ,51(8), 2043-2053.

[8] E. Lawrence, G. Michailidis, V. N. Nair. Statistical Inverse Problems in Active Network Tomography, *IMS Lecture Notes - Monograph Series*,2007 ,Vol(54), 24-44.
[9] P. McCullagh and J. Nelder. Generalized Linear Models, *Chapman and Hall*,1989.
[10] X. L. Meng, D. A. Dyk. The EM algorithm- An old Folk Song Sung to a Fast New Tune, *J. R. Statist. Soc, Ser.B*,1997 ,59 ,511-567.
[11] C. F. J. Wu. On the Convergence Properties of the EM Algorithm, *Annals of Statistics*,1983 ,11, 95-103.