

A Framework of Video Network Measurement Based on Utility Optimization

Yu Zhao^{1, a}, Bo Fu^{1, b} and Kai Cao^{1, c}

¹Zhuhai Power Supply Bureau of Guangdong Power Grid Corporation, Zhuhai City, Guangdong Province, 519000

^a13825603652@139.com, ^b446137166@QQ.com, ^c249467404@QQ.com

Key words: the packet level measurement; network measurement; utility optimization; detection rate

Abstract: In recent years, the packet level measurement is widely used to evaluate the packet loss and delay performance in broadband networks. The samples of packet loss and delay are provided by probe packet in active probe, from these samples can calculate the overall performance of the communication. However, this method will bring some errors: such as the error caused by insufficient sampling, self-interference generated from too many probe packet injection, and the deviation from the sample. Therefore, this paper presents an analytical technique to quantify the overall utility function associated with the following factors: (i) disturbance caused by each probe packet, (ii) deviation and (iii) the variance of the probe (sampling) function. The simulation results show that the measurement of the optimal detection rate depends on the type of parameters used by the network engineers.

Introduction

In recent years, the measurement results of packet networks are more and more widely used in the Measurement-Based Admission Control. In the MBAC program, the status of the network must to be quickly accessed and the decision on whether to accept a new connection must to be make immediately. Therefore, there are more and more research on the measurement of packet networks.

In the literature [1], the end to end application based on service quality is discussed, and a new tool for network measurement is proposed. Wireless MBAC system is proposed in the literature [2]. By studying the statistical inference in the queue, Clarke [3] derived the maximum likelihood estimator (MLE) in the M/M/1 queue. By comparing the relative efficiency of the direct estimation of the average waiting time, Jenkins believes that MLE has a low gradual change, especially for high load values. In 1974, through summarize and compare the estimating functions of different arrival and departure rates in the M/M/1 queue, Aigner[4] conclude that, even for simple settings, the estimating function is still a large number; subsequently, Aigner uses a different valuation as the standard to find the best asymptotic variance. However, this does not apply to a small sample inference, but also did not consider the time required to collect a large amount of data. Through the study of the covariance structure of the queue, Reynolds[5] assesse the different estimates of variance. In the case of using different sampling frames, Reynolds according to the fixed time to observe the queue, while the Aigner is in a fixed number of customers to observe. Basawa and Prahuprove that MLEs is asymptotically normal distribution, and give an example of application in M/M/1 queue. This work utilizes the probabilistic results in Billingsley. After that, Basawa and other set up a general framework for finding the Fisher information matrix, which has great use value for the calculation design. This method is also used in the literature [6].

Previous studies have shown that the traditional evaluation function is effective in the case of long time observation of queues. However, the research on how to measure the queue in a relatively short time is less. Therefore, the purpose of this study is to find a method to measure the packet network in the shortest possible time.

In term of statistics, our concern is to find the estimating function to measure certain properties of the network. An estimating function receives some observable data and generates an estimated

value of the unknown parameters which we are interested in. Estimated value is a specific value of an estimating function under given specific data. We realize the optimization from the following three aspects: (1) Estimated bias; (2) Estimated variance; (3) Disturbance caused by active detection.

The structure of the remaining part is as follows: in the second section, the theoretical basis of data packet network measurement is introduced in detail, and a general utility function is proposed. In the third section, the application of the utility function method is introduced, and the application example of the single M / M / 1 queue is mainly introduced. Simulation experiments and analysis are carried out in the fourth section. Conclusion is made in the fifth section.

The theoretical basis of data packet network measurement

In this paper, we consider all the network measurements as a numerical test of sampling using random processes, and evaluate the validity of sampling using the bias and variance of the estimating function. In this way, we can use the statistical principle of DOE to carry out the experiment of network measurement, so as to develop a method to find the best active detection rate. Therefore, we propose a kind of utility function which can synthetically consider the deviation and variance of the estimating function and the network disturbance caused by the detection.

Symbolic model overview. Assume that the data packet service arrive at the network with a fixed rate λ , and receive the service and leave the network at a fixed rate μ .

We want to find the value of some unknown parameters in the network θ , such as the packet loss rate and average system delay for interfering traffic packets. The value of these unknown parameters in network θ cannot be directly measured, but it can be detected when the data appears after the introduction of probe data packet at rate x . We can deduce the network θ according to the time of the probe data packet in the system. Therefore, the goal of this paper is to find the x value that can correctly estimate the network θ .

In this paper, $S(t)$ is used to represent the time required to all data packets complete the service and exit the network in time t network, which we call the virtual waiting time.

Here, $S(t)$ is a non-negative right continuous time process. When a packet enters the system at the time a_0, a_1, a_2, \dots , the instantaneous jump occurs, the jump amplitude is determined by the queuing rule, and is consistent with the duration of the packet. For example, the magnitude of the jump in the M / M / 1 queue is consistent with the corresponding service time, and follows the exponential distribution of the parameter μ .

Jump time a_i and jump amplitude $S(t)$ are random variables. Unless we have full understanding for the queue, we cannot directly observe $S(t)$. But we can speculate by introducing a probe packet at time $T_1, T_2, T_3, \dots, T_N$. Through introducing new data packet into the system, a new process $S^*(t)$ is formed, as shown in Figure 1.

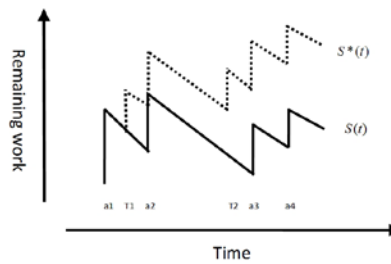


Figure 1. $S(t)$ and $S^*(t)$

The number of probes generated during the study was above N (Note that generally are random variables, but we can make it fixed. For example, by setting the time interval between probes to a fixed value). We recorded the time between the probe packet j entering and leaving the system

as y_j , that is, the system time. Our concern is the estimated value $\hat{\lambda}$ of λ . Given $N=n$ groups, we have $t_1 < t_2 < \dots < t_n$, and observe $S^*(t_1), S^*(t_2) \dots S^*(t_n)$ under no error. Let $y_j = S^*(t_j)$, data can be represented by $y_1, y_2 \dots y_n$, writing as vector \mathbf{y} .

Utility function. Since the above analysis method is not easy to deal with the general network, we use a utility based approach to optimize the detection rate. We must first consider what we need to achieve when we measure the network. Obviously, we want to accurately measure the unknown parameters of the network (such as the average delay, packet loss probability, etc.). In other words, we would like to minimize the bias of the estimated value $\hat{\theta}$ of the network θ , which is derived from the function of \mathbf{y}, \mathbf{x} and μ . The bias of the estimated value is the difference between the expected value and the true value of the estimated value, that is, $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$. By the statistical theory of the experimental design, it is known that it is desirable to minimize the variance $\hat{\theta}$ of the estimator. Ideally we would want a lower bias and variance of the estimator. We use the following formula to measure each data packet of detection rate x :

$$D(x) = \frac{1}{N} \left(\sum_{i=1}^N [c(S^*(a_i) - S(a_i))] \right)^{1/r} \quad (1)$$

$c(w)$ is a cost function of data packet delay w , $r > 0$. Generally, the underlying $S(a_i)$ is not directly observable, and we estimate it through simulations. The bias of the exponential r 's nonlinear penalty and average time delay, such as in VoIP system, a data packet with no delay and a data packet with a 20ms delay are more destructive than the two data packets with 10ms delay.

In order to minimize the bias, variance and disturbance, we propose a general utility function:

$$\psi(Bias(\hat{\lambda} | x), Var(\hat{\lambda} | x), D(x)) \quad (2)$$

And use it to find out the optimal detection rate

The specific form of the utility function will depend on how we balance accuracy and the disturbance caused by detecting at this rate. When the bias and variance are combined, a natural measure is the mean square error.

$$MSE(\hat{\lambda}) = [Bias(\hat{\lambda})]^2 + Var(\hat{\lambda}) \quad (3)$$

In order to balance variance and disturbance, we propose a general form of utility function:

$$\psi(x) = -[\kappa \sqrt{MSE(\hat{\lambda})} + (1 - \kappa) D(x)] \quad (4)$$

In which $D(x)$ has been defined in (1), $0 < \kappa < 1$. The selection of κ is important, it balances the importance of measurement and the other costs caused by the disturbance. If we don't care about the disturbance caused by the measurement, we choose $\kappa = 1$; If we don't care about measuring accuracy, select $\kappa = 0$. For most practical applications, we will choose κ between 0 and 1.

We present the utility function without units. In a sense, it is a mapping from two-dimensional domain to one dimension range (utility). This utility can be considered a two-dimensional input of the sorting mode.

The basic application of data packet network measurement

We tried to simulate a real active detecting environment, in this environment, we may need to quickly estimate. For example, In MBAC, It can be assumed that the network flow rate λ is constant only in a short period of time, and in most cases we want to quickly find a valuation of the λ . So we assume that to perform each experiment in a short period of time T , and then we perform m experiments with different values of x to estimate the effectiveness of the

experimental implementation.

For the network we studied, based on fixed service rate μ , arrival rate λ and other parameters can determine the user traffic, and based on the current network conditions determine the possible detection rate set \mathcal{X} .

Basic steps of network measurement. The basic steps of network measurement are as follows:

- 1) Choose a detection rate $x \in \mathcal{X}$;
- 2) Simulation of a queue without detection;
- 3) Continue to simulate the queue, introduction of the detection data packet by ratio x , to find the time consumed by the detection packet in the system;
- 4) When the simulation queue has been carried out to our choice of experimental time T , we use the selected estimator to form an estimated value of $\hat{\theta}$;
- 5) To re-run the simulation by using random number generated by the potential traffic arrival and service time, in the absence of detection to assess the time delay due to the detection;
- 6) Repeat steps 2)~5) to estimate the mean deviation and variance of $\hat{\theta}$, and calculate the average packet delay in the potential queue;
- 7) Repeat steps 2)~6) to assess bias and variance of $\hat{\lambda}$ for different values of x .

Single M/M/1 queue. Now select a simple queue for improvement, initially assuming that the network consists of a single M / M / 1 queue. Although the network is not representative, but in order to clearly demonstrate our approach, here to present a simple example for analysis.

For the M/M/1 queue, we perform the service at the rate of $\mu=5$ per second, assuming that we want to estimate the arrival rate θ . Any reasonable estimator may be selected, and different estimators have different estimates in general, so there are different optimal ratios. We choose an estimator based on Aigner.

$$\hat{\theta} = \frac{(1/N) \sum_{i=1}^N Y_i - (1/\mu)}{(1/N) \sum_{i=1}^N Y_i (1/\mu)} \quad (5)$$

This estimation function is only to demonstrate the effectiveness of the method to determine the optimal detection rate, not to assess the quality of the estimated function.

Simulation results and analysis

Simulation setting. We set $\lambda = 2.5s^{-1}$. Readers see here $A = B$, it is a simulation parameters we want to measure, but when we use (5) to estimate $\hat{\theta}$, cannot use the knowledge of this λ .

We make the candidate x in 0.1s to 2.4s with interval change 0.1, and will be strictly limited to scope of $\lambda + x < \mu$. We set a 10s pre simulation run period, after which assume that can perform the detection experiment of $T=10s$. For each candidate point, we perform $m=1000$ simulation, using $1000 \hat{\lambda}$ generated by each candidate point to estimate bias and variance of $\hat{\lambda}$. And get the underlying virtual system time process $S(t)$ of simulation and process $S^*(t)$ which changed after detection. We can estimate the value of perturbation function $D(x)$, let $c(z) = z$. We penalized every "linear" packet delay. In order to elaborate a possible utility function, we set up $\kappa = 1/2$.

Simulation result.

Detection result. The detection results are shown in Figure 2, the optimal detection rate in the right image below, which $x \approx 1.2s^{-1}$

In the Y axis of the logarithmic scale, we see that at low rate obtained significantly worse results. As shown in Figure 2 on the top left corner, poor results at lower x are mainly due to the very high deviation estimates for the lower x . Because we restrict the experiment to the 10s, the number of the experimental Poisson probes for the low rate is very small. For example, $x=0.1$, we can only expect a probe packet, so the result is poor.

As expected, the deviation of estimator $\hat{\lambda}$ will occur both at a larger or a smaller detection rate x . The increase of deviation tells us the main problem of active detection is self-interference.

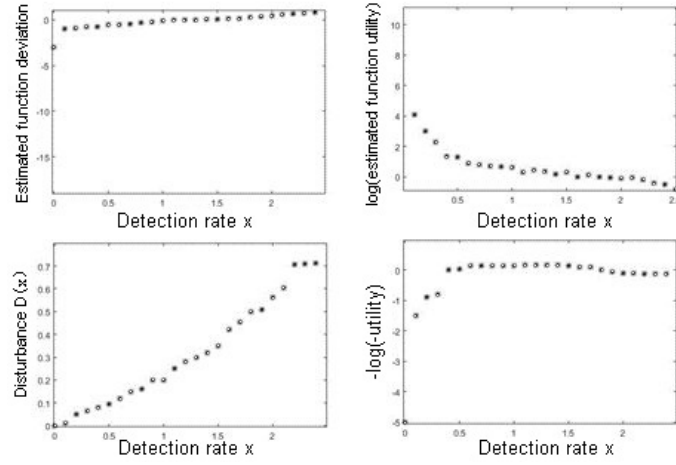


Figure 2. The bias, variance, disturbance and utility of the utility function of the M/M/1 queue with the change of the detection rate X

The influence of different parameters on utility function. the utility function changes with r .

Figure 3 shows the effect of utility change with r , we only focus on the high detection rate range ($x > 0.6$). By finding the optimal detection rate through the discovery of high efficiency in this interval, we find that the R value does not play a significant role. From the simulated data, the optimal detection rate is $x = 1.2s^{-1}$. The absolute value of the utility (i.e., the height of the graph) is not important, because we are only seeking the optimal detection rate for a fixed r .

Figure 4 shows the variation of the utility when we set the $\kappa = 0.05$. Changes in the value of κ may represent the network engineer are now more concerned with minimizing the network delay rather than reducing the mean square error of our estimator. We see here, the best detection rate is lower, for all r , the three sets of data obtained maximum value in $x \approx 0.5s^{-1}$.

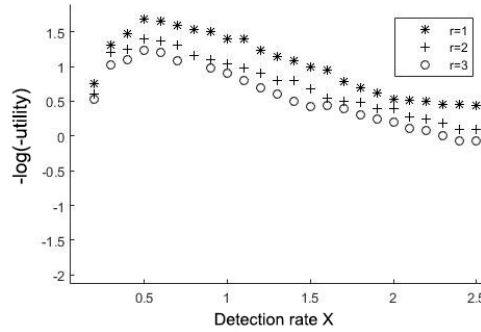


Figure 3. when $\kappa = 0.05$, r taken 1, 2, 3, the utility changes with the change of detection rate

Practical problem analysis.

The deviation and disturbances detected in a network off-line. We have demonstrated that, if fully understood a particular network, we can determine the optimal detection rate. In practice, we may not fully understand network, in particular, when the utility function is established, we consider that the deviation and perturbation function are known, but in practice we cannot determine.

We can use the simulation method to estimate the deviation and the disturbance, establish a simulation network offline and use the technology in this paper to determine the optimal detection rate. For the simulator, the response of the network can be observed in the presence or absence of active detection, to evaluate the function of the deviation and perturbation function, from which we can find an optimal detection rate. Although our optimal detection rate is accurate for the simulation

model, the difference between the model and the actual network usually means that the optimal rate of a model is second best to the other one. It is not uncommon to approach complex networks through a simple network, plan to assess how the difference between the actual network and the simulator can affect the success rate of the measurement in the future work.

Assumptions applied in the sample. In our case, we make some assumptions: stability of the network, that is, the probability of arrival of traffic is constant with time; Simplifying network properties such as first come to first serve; Lack of network routing. In this paper, the method for finding the optimal detection rate is the same as in the case of whether these assumptions are used, but if we relax these assumptions, it may lead to bias, disturbance, especially the variance will be a great change. These networks are selected, because they are simple networks that we are familiar with, and can easily be simulated in order to estimate the disturbance and the deviation.

Conclusion

In the literature [4], under the condition of considering the active detection cause interference, demonstrates how to approximate a queue through the Markov chain model, and found a precise optimization design model. In this paper, we obtain the optimal design (optimal detection rate) by estimating the deviation, variance and disturbance in the simulation.

This approach is very flexible and does not require a lot of understanding of network, or you can use any of the theoretical queuing models to determine an optimal detection rate. We only need to quantify the important information in the measurement, for example, at the cost of network disturbance, we hope that the network measurement is more accurate. Through simulation, the results show that under the assumption of any network can find the optimal rate. This is a valuable tool for network practitioners when considering the purpose of successful measurement. In fact, in many network monitoring applications, this tool also has a great practical value.

Acknowledgements

The research and development applications of video monitoring quality diagnosis system in substation (GDDW1020150301DK00002).

Reference

1. Chen, L., Sun, T., Yang, G., Sanadidi, M., Gerla, M.: Ad hoc probe: path capacity probing in wireless ad hoc networks[C]. First Int. Conf. On Wireless Internet, 2005. Proc., 2005, pp. 156 – 163.
2. Su, Y., Wu, R., Li, R., Duan, W., Luo, J.: A novel measurement-based call admission control algorithm for wireless mobility networks under practical mobility model[C]. 2010 Int. Conf. on Communications and Mobile Computing, IEEE, 2010, pp. 364 – 368.
3. Na, Y., Jun, S., Yuxin, C., Ye, J.: Measurement-based admission control algorithm for wireless IP network[C]. WRI Int. Conf. On Communications and Mobile Computing, 2009. CMC' 09, 2009, pp.216-220.
4. Parker, B.M., Gilmour, S.G., Schormans, J.: Measurement of packet loss probability by optimal design of packet probing experiments[C]. IET Commun. Special Edition on Simulation, Analysis and Measurement of Broadband Network Traffic, 2009, 3, (6), pp. 979 – 991. URL <http://www.DigitalObjectIdentifier10.1049/iet-com.2008.0075>
5. Clarke, A.: Maximum likelihood estimates in a simple queue, Ann. Math. Stat., 1957, 28, (4), pp. 1036 – 1040.
6. Jenkins, J.: The relative efficiency of direct and maximum likelihood estimates of mean waiting time in the simple queue M/M/1[J], J. Appl. Probab., 1972, 9, (2), pp. 396 – 403.