Loss Analysis for Networks based on Heavy-Tailed and Self-Similar Traffic

Danna Zhuang and Chuanhuang Li*
School of Information and Electrical Engineering, Zhejiang Gongshang University, Hangzhou, China
*Email: chuanhuang_li@mail.zjgsu.edu.cn

Abstract. Many businesses on the computer network appear heavy-tailed self-similarity (long range dependency), which means network traffic exists burst. Various service source with burst characteristic which show self-similar have a significant effect on transmission performance, network traffic control strategy and network performance indicators such as loss. Loss is an important QoS parameter at the network node, which need to be considered and controlled in all types of traffic, but there is no paper study loss analysis based on heavy-tailed self-similarity. Thus, only analyze and evaluate loss under self-similar traffic can reduce the adverse effects which caused by traffic self-similarity and optimize network performance. We adopt stochastic network calculus approach to rewrite loss issue. Based on this, we present the loss analysis in the case of a single node network using heavy-tailed service curve and heavy-tailed self-similar arrival curve, and loss analysis under cross traffic, multi-node networks with concatenation. And we also get the relational graph between the loss and the arrival rate, service rate.

1. Introduction
The Bellcore laboratory[1] demonstrated the network traffic exhibits self-similarity for the first time, the phenomenon of self-similar truly reflects the network flow will occur burst after a long time, which means that the network presents long range dependency. Then people gradually find that heavy-tailed[2] files sizes and bursts is the main reason to form self-similar and long range dependency phenomena, which means that the portion of heavy tail resulted in the occurrence of self-similar.

To monitor and analyze a lot of traffic (such as WAN, LAN, VBR and ISDN, etc.), the results have shown that a variety of business on the computer network appear statistical self-similarity (long range dependency), which means network traffic exists burst. So the tradition traffic model is no longer suitable for the analysis of current network traffic. Self-similar model has become a hot topic for the researchers, because it can describe the characteristics of network traffic more accurately.

Existing research on stochastic service guarantees focuses on two aspects: latency and backlog. Loss is an important QoS parameter at the network node, which need to be considered and controlled in all types of traffic. Loss of performance analysis is applied to multiple scenarios, such as connection admission control and packet scheduling algorithms. In many cases, the results acquired from the buffer overflow probability are used to depict the loss approximately [3,4,5,6]. Yet, this approximation basically provides loose bounds.

To evaluate the performance of computer networks, Stochastic network calculus[7] is usually used. But so far, for loss analysis, it is not common, compared with deterministic network calculus. Of
course, we can also find some theories in [8, 9]. But the upper limit of the loss rate is only obtained by one particular buffer size like b. In[10], the loss factor is proposed as a new parameter, based on this, the loss bounds are calculated from existing arrival curve and service curve. But the loss algorithm in this article also has more loopholes, we will briefly illustrate these vulnerabilities or errors in section III.

The paper[11] proposed a non-asymptotic delay analysis of a multi-node networks with heavy-tailed self similar traffic and heavy-tailed service. But there is no paper study loss analysis. So it is significant to pay attention to the loss behavior and make the results according to the theory of heavy-tailed self-similar and stochastic network calculus better.

The remaining sections are structured as follows. We will discuss following content in section II: (1) the notations of stochastic network calculus used are introduced, (2) traffic model based on statistics and especially heavy-tailed self-similar traffic envelope process are introduced. (3) the characterization of heavy-tailed as well as self-similarity. Then we presents the loss analysis of heavy-tailed service and heavy-tailed self-similar traffic networks, and loss analysis under cross traffic, multi-node networks with concatenation in section III, also the loopholes of loss algorithm in[10] also will be illustrated. The relational graph between the loss and the arrival rate, service rate will be introduced in section IV.

2. Theoretical Background

2.1. Notation

The notations used in this paper are defined as follows: $A(s, t)$ indicates the number of bits from a flow to a node. $D(s, t)$ indicates the output flow from the node within the interval $(s, t]$. $S(t)$ indicates the number of service (in number of bits) provided by the network element, $L(t)$ indicates loss process in time interval $(0, t]$. For any $0 \leq s \leq t$, there are $A(s, t) = A(t) - A(s)$, $D(s, t) = D(t) - D(s)$ and also $S(s, t) = S(t) - S(s)$, $L(s, t) = L(t) - L(s)$.

Let $f \otimes g(t) = \inf_{s \in \mathbb{R}} \{f(t - s) + g(s)\}$ represent the minimum plus convolution of functions $f$ and $g$, and accordingly, “infimum” should be construed as “minimum”.

Let $f \oslash g(t) = \sup_{u \in \mathbb{R}} \{f(t + u) - g(u)\}$ represent the minimum plus deconvolution of functions $f$ and $g$, “supremum” should be construed as “maximum”.

2.2. Heavy-tailed self-similarity

Self-similarity refers to the partial structure and the overall structure has a certain degree of consistency if they compared, i.e., observation of an object at different spatial dimensions, which have a similar morphology. Network traffic which has self-similarity showing sub exponential decay when the queue is long. The important parameter of self-similar characteristic is Hurst parameter, the bigger Hurst parameter, the higher self similar process, and the bigger burst.

[11] shows the definitions of heavy-tailed and self-similar. If the tail distribution of the random process $X$ satisfies the law of power $P(X(t) > x) \sim Kx^{-\alpha}$, then we say that it has a heavy-tailed distribution, and it has a tail index $\alpha \in (0, 2)$ and a scaling constant $K$. If the tail index in $1 < \alpha < 2$, we say that the mean of the distribution is limited, but the variance is infinite. If the whole and part of a random process, one part and other parts of the complex process have the same distribution in the fine structure or property, we say that the random process is content to the self - similarity. For each $\alpha > 0$, there is $X(t) \sim_{\alpha} \alpha^{-H}X(\alpha t)$. Where the index $H \in (0, 1)$, called the Hurst parameter, is used to specify the extent of self-similarity. If a process is satisfies two conditions above at the same time, we will call this process as heavy-tailed self-similar process.

2.3. (Htss) envelope process and (ht) service curve based on statistical traffic model

Statistical traffic model is different from other traffic model. It cast aside specific morphological
characteristics of network traffic, use the statistical regularity restrictions of the network traffic to restrict directly and study the performance of network which has the traffic restriction characteristics. Statistical traffic model gives traffic process upper bound of data stream which use the probability significance. It mainly adopt statistical envelope[11] and statistical sample path envelope to describe the traffic.

Given an arrival process $A(t)$, which is cumulative in the time interval $[0,t]$, function $\zeta(t-s;\sigma)$ is non-negative and non-decreasing about $t-s$ and $\sigma$, if for all $s,t \geq 0$ and for all $\sigma > 0$, that

$$P(A(s,t) > \zeta(t-s;\sigma)) \leq \epsilon(\sigma)$$

We say $\zeta$ is the statistical envelope of the arrival process $A(t)$, where $\epsilon(\sigma)$ is a non-increasing function about $\sigma$ , when $\sigma \to \infty$, $\epsilon(\sigma) \to 0$, we call $\epsilon(\sigma)$ as violation probability.

When deriving the traffic envelope, usually write statistical envelope process in the flowing form: $\zeta(t-s;\sigma) = \rho(t-s) + \Delta(t-s;\sigma)$, where $\rho(t-s)$ corresponds to the mean rate of traffic. And $\Delta(t-s;\sigma)$ is burst size, which is non-negative function. There are two forms, one is statistical envelope with burst size $\sigma$ and the other one is burst size with function. This article we analyzed is one functional form: heavy-tail self-similar (htss) envelope process. Our paper adopted stochastic network calculus approach to rewrite loss issue. We describe the traffic as statistical envelope, which define upper bounds of traffic in the interval of time. The service flow is modeled by service curves that define a lower bound on the available services for the flow.

Liebeher[11] presented envelope processes which are provided with heavy-tail self-similar characteristic. If arrival processes satisfy heavy-tail self-similar (htss), it can be written as

$$P(A(s,t) > \rho(t-s) + \sigma(t-s)^H) \leq K\sigma^{-\alpha}$$

$K$ and $\rho$ are constants. $H$ is Hurst parameter, and $H \in (0,1)$, $\alpha$ is tail index, and $\alpha \in (0,2)$, where the mean of arrivals is finite mean but the variance of it infinite. It is a statistical envelope

$$\zeta(t;\sigma) = \rho t + \sigma^H, \epsilon(\sigma) = K\sigma^{-\alpha}$$

The probability of straying from the average rate $\rho$ submit to the law of power in the htss envelope. The probability of deviation is likely to increase over time due to self similarity. The htss envelope specifies boundaries to describe various traffic, but unless this traffic has heavy-tailed self-similarity, the performance will not be loose.

[11] define the heavy-tailed (ht) service curve as the following form of service curve

$$S(t,\sigma) = [Rt - \sigma], \epsilon(\sigma) = L\sigma^{-\beta}$$

$L$ is constant, and $0 < \beta < 2$. The deviation from the service rate guarantee $R$ which specifies by the ht service curve has a heavy-tailed decay. We can see that the definition in formula (4) of the HT service guarantee does not include Hurst parameter, the reasons for this is beneficial to the calculation of the service bounds at more than one node.

3. Loss analysis based on heavy-tailed and self-similar

Mathematical model which based on the self-similar traffic has become indispensable realization elements of performance evaluation and optimization, traffic control and network construction flow for the current network. And it also have the important theoretical and practical application value on the setting of network planning, network control, and high-quality network services. Various service source with burst characteristic which show self-similar have a significant affect on transmission performance and network traffic control strategy. And it also has the direct impact on network performance indicators such as loss, making network design, control, analysis and management become complicated. Thus, only analyze and evaluate loss under self-similar traffic can reduce the
adverse effects which caused by traffic self-similarity and optimize network performance. In this section, we firstly give the definition of loss period, then present the loss bound based on heavy-tailed and self-similar.

### 3.1. Single node loss analysis

Loss period\([10]\): Generally, in a network system, there is an arrival rate and a service rate. When the buffer is full and the former is greater than the latter, we call this period of time as the lost period. In that article the author divided the lost period into many small loss periods, but it can not be divided according to the definition. We assume that \((s, t]\) is the lost period, in this period the amount of loss is equal to the amount of arrived subtract the amount of service. So

\[
L(s,t) = A(s,t) - S(s,t)
\]

(5)

There is a sentence written that in order to avoid a system crash, the arrival traffic should not exceed the service provided by the network system in the long term \([9]\). This sentence has a problem, because we consider that in the lost period, arrival traffic certainly exceed the service. So assumption \(x > 0, \beta(x) > r(x)\) in \([9]\) is erroneous.

Theorem 1: When the buffer of the network system is limited, there is an arrival traffic. This flow is described by an hts envelope \(P(A(s,t) > r(t-s) + \sigma(t-s)^H) \leq K\sigma^{-\beta}\), and the system provides the h service curve at a node given by \(S(t, \sigma) = \{ R(t - \sigma) \}_\sigma, \epsilon(\sigma) = T\sigma^{-\beta}\), for all \(l > 0\) and \(0 \leq s \leq t\), that

\[
P\{L(s,t) > \ell\} \leq \inf_{0 \leq \ell < \ell} \left\{ \int (l - r \cap R(0)-t)^\beta + K t^{-\ell(1-H)} \right\}
\]

(6)

Proof: From (5), we can know that

\[
L(s,t) = A(s,t) - S(s,t) = A(s,t) - r(s,t) + r(s,t) + R(s,t) - S(s,t) = A(s,t) - r(s,t) + R(s,t) - S(s,t) + r(s,t) - R(s,t) \leq \sup_{0 \leq s \leq t} \{ A(s,t) - r(s,t) \} + R(s,t) - S(s,t) + \sup_{0 \leq t \leq l} \{ r(t) - R(t) \}
\]

(7)

We can get \(P\{L(s,t) > \ell\}\) by the right part of (6); that is, \(P\{ \sup_{0 \leq s \leq t} \{ A(s,t) - r(s,t) \} > \ell\}\) and \(P\{R(s,t) - S(s,t) > \ell\}\) and \(\lim_{\ell \to +0} \frac{1}{\ell} [ r(t) - R(t) ] \leq 0 \).

Because the flow is described by an hts envelope with \(\zeta(t; \sigma) = r(t) + \sigma^H, \epsilon(\sigma) = K\sigma^{-\beta}\), according to formulas \(S(t, \sigma) = \{ R(t - \sigma) \}_\sigma, \epsilon(\sigma) = T\sigma^{-\beta}\), we can get the h service curve of a node. From (2) we can derive that \(P\{ \sup_{0 \leq s \leq t} \{ A(s,t) - r(s,t) \} > \ell\} \leq K t^{-\ell(1-H)}\); \(P\{R(s,t) - S(s,t) > \ell\} \leq T t^{-\beta}\).

We can conclude immediately from (6)

\[
P\{L(s,t) > \ell\} \leq K t^{-\ell(1-H)} \cap T(l - r \cap R(0)-t)^\beta \leq \inf_{0 \leq \ell \leq \ell} \left\{ \int (l - r \cap R(0)-t)^\beta + K t^{-\ell(1-H)} \right\}
\]

(8)

Proof end.

### 3.2. Loss analyze under cross traffic

We call the traffic which should be analyzed as through traffic, while call the other traffic as cross traffic, and through traffic is guaranteed under the premise of the cross traffic to be served first.

We consider the case where through traffic and cross traffic compete because of the resources in a system which at a constant-rate link with a capacity of C, as shown in figure 1. Let A and Ac
respectively be the arrival of through traffic and cross traffic. Let \( A_c \) be described by an hts envelope as \( \zeta_c(t) = r_c(t-s) + \sigma_1(t-s)^\alpha \), \( \varepsilon(\sigma) = K_1\sigma^{-\alpha} \) where the arrival rate bound in accordance with \( r_c < C \). Let A be described by an hts envelope as follows \( \zeta(t) = r(t-s) + \sigma(t-s)^\alpha \), \( \varepsilon(\sigma) = K\sigma^{-\alpha} \), the through traffic can be guaranteed by a hts left over service curve \( Q = S(t; \sigma) = [(C - r - \mu)t - \sigma], K = \varepsilon(\sigma) = T\sigma^{-\beta} \) where \( T \) is constant, \( r_c \) means the velocity of cross traffic, \( \mu > 0 \) is a free parameter. Let D and Dc respectively be the arrivals departures of through traffic and cross traffic. According to the statistical envelope of self-similar heavy-tailed and left over service curve, that we can derive a self-similar heavy-tailed through traffic while there are also exit cross traffic the worst-case performance provided by service the system.

\[
\begin{align*}
\text{Figure 1. A system that competes for resources.} \\
\end{align*}
\]

So the (6) we can write

\[
P[L(s, t) > l] \leq KL^{-(l-H)} \otimes T(l - r \oplus (C - r - \mu)(0))^{-\beta} \leq \inf_{\alpha, \beta} \left[ T(l - r \oplus (C - r - \mu)(0) - t)^{-\beta} + K(l - r(t - l)) \right] (9)
\]

3.3. Loss analyze under concatenation

The concatenation property[11]: When a series of servers are in a concatenated situation, we can treat them as one server and use the same server model to represent.

As shown in figure 2. For all nodes \( n = 1,2,..N \) concatenated in tandem, there is an ht service curve \( S_n(t, \sigma) = [Rt - \sigma], \varepsilon(\sigma) = L\sigma^{-\beta} \) which guarantees the service. Then an arrival flow traverses these nodes.

\[
\begin{align*}
\text{Figure 2. Concatenation network system} \\
\end{align*}
\]

For \( \theta > 1 \), the entire concatenation network system provide service guarantee is

\[
S_{net}(t, \sigma) = [R(t - \sigma)], \\
\varepsilon_{net}(\sigma) \leq N^{2\beta} \cdot 2^{\beta-1} \cdot \varepsilon(\sigma) \left| \log \varepsilon(\sigma) \right| + (1 + \beta) \log N + 2 \quad (10)
\]

where: \( \varepsilon(\sigma) = \min \left\{ 1, \frac{2^{\max(I,\beta)}}{\beta \log \theta} L\sigma^{-\beta} \right\} \).

According to loss analysis in a single node and cross traffic, we can derive loss analysis on the concatenation system easily.

\[
P[L(s, t) > l] \leq KL^{-(l-H)} \otimes \varepsilon_{net}(l - r \oplus (R/\theta)(0)) \quad (11)
\]
4. Numerical Results

Next, we mainly discuss the loss algorithm derived in section III, and get the relational graph between the loss and the arrival rate, service rate. As we mentioned earlier the tail index $\alpha \in (0, 2)$ and if the tail index in the range $1 < \alpha < 2$, we say the distribution has a limited mean, but the variance is infinite. And the Hurst parameter $H \in (0, 1)$, $0 < \beta < 2$, so the parameters are $\alpha = 1.98$, $H = 0.93$, $\beta = 1.2$. The remaining parameter K and T we need to know in the above formulas can be obtained by taking the (2)[11], we choose $K=1225$, $T=1000$. We use LINGO[12] to analyze the formula (7), and let $l=100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800$, first let $r=10, 20, 40, 60, 80, 100$(Mbps). And use the OriginPro[13] convert these data into graph.

We can get some laws from figure 3(a). One is the $P_{L(s,t)>l}$ reduces as the $l$ (the amount of loss) increases. Second, the bigger $r$, the bigger $P_{L(s,t)>l}$, which means that the bigger arrival rate, the bigger $P_{L(s,t)>l}$. From the $(l-r \oplus R(0))$ in formula (7), we know that $r \oplus R(0) = \sup_{t \geq 0}[r(t) - R(t)]$, so the relationship between arrival rate and loss $P_{L(s,t)>l}$ is opposite to the relationship between service rate and loss $P_{L(s,t)>l}$ as shown in figure 3(b). The bigger service rate, the smaller $P_{L(s,t)>l}$.

![Figure 3](image.png)

**Figure 3.** The relational graph between the loss and the arrival rate, service rate.

5. Conclusion

Research shows that variety of business on the computer network appear statistical heavy-tailed self-similarity, and loss is an important parameter of QoS. So it is great interest to study the loss behavior under the heavy-tailed self-similar. We perform a loss analysis on traffic characterized by hsss envelopes, and receive services provided by heavy-tailed service curves. We also get the relational graph between the loss and the arrival rate, service rate. The next stage of work will be to conduct our research on other arrival curve and service curve.

Acknowledgment

This work was supported in part by a grant from the National Key Research and Development Program of China (No.2017YFB0803202), the Natural Science Foundation of China(61871468), the Zhejiang Provincial Natural Science Foundation of China(LY18F010006), the Zhejiang’s Key Project of Research and Development Plan(2017C03058), the Zhejiang Provincial Key Laboratory of New Network Standards and Technologies (No.2013E10012), Scientific Research Projects of Zhejiang Education Department(Y201737967).
References

[1] Leland W E, Taqqu M S, Willinger W, and Wilson D V 1994 On the self-similar nature of Ethernet traffic IEEE/ACM Transactions on Networking 2(1) 1–15.

[2] Robert J A, Feldman R, Taqqu M S 1998 A practical guide to heavy tails : statistical techniques and applications Boston : Chapman & Hall.

[3] Rocha F G C and Vieira F H T 2018 A Channel and Queue-Aware Scheduling for the LTE Downlink Based on Service Curve and Buffer Overflow Probability IEEE Wireless Communications Letters 8(3) 729-732.

[4] Ciucu G, Poloczek F, and Rizk A 2019 Queue and Loss Distributions in Finite-Buffer Queues in Proceedings of the ACM on Measurement and Analysis of Computing Systems New York, NY, USA, June, doi: 10.1145/3341617.3326146.

[5] Chydzinski A, Samociuk D, and Adamczyk B 2018 Burst ratio in the finite-buffer queue with batch Poisson arrivals Applied Mathematics and Computation 330 225-238.

[6] Wahid A A, Schormans J A and Wahid A A 2015 Bounds on accuracy when estimating the loss probability in a packet buffer in Proceedings of 2015 7th Computer Science and Electronic Engineering Conference (CEEC) Colchester, UK.

[7] Jiang Y and Liu Y 2008 Stochastic Network Calculus (Berlin/Germany Springer).

[8] Andrá s Gulyá s and Jó zsef Biró 2006 A stochastic extension of network calculus for workload loss examinations IEEE Communications Letters 10(5) 399–401.

[9] Gulyá s A, Biró J, and Né meth F 2009 New stochastic network calculus for loss analysis in Proceedings of the 4th International ICST Conference on Performance Evaluation Methodologies and Tools (Valuetools ’09), ICST(Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Brussels, Belgium, vol.38 pp 1–10.

[10] Luo C, Yu L, and Zheng J 2013 Extending Stochastic Network Calculus to Loss Analysis, ”The Scientific World Journal Article ID 918565, 8pages.

[11] Liebeherr J, Burchard A and Ciucu F 2009 Delay bounds for networks with heavy-tailed and self-similar traffic Technical Report, <arXiv:0911.3856v1>.

[12] CHILDRESS, John 2014 Learning the lingo Economist vol. 410 Issue 8869 p72-73. 2p.1 Color Photograph.

[13] Seifert E 2014 OriginPro 9.1: Scientific Data Analysis and Graphing Software—Software Review Journal of Chemical Information and Modeling 54(5) 1552–1552.