

Study of Network Traffic Analysis Model

Xiaoling Tan and Zefu Tan
College of Physics and Electronic Engineering
Chongqing Three Gorges University
Chongqing, China
fenglingcao86@163.com

Weijian Fang
Experiment Center
Chongqing Three Gorges University
Chongqing, China
wanzhoufwj@yahoo.cn

Abstract—Network traffic analysis model is the base of traffic predication and network evaluation. The model is improving when the knowledge about the traffic features being abundant. According to the different features, the paper introduced the typical models in modern network application, and studied on their characteristics.

Keywords— Traffic analysis model; Self-similarity; Dependency

I. INTRODUCTION

As the fast increasing of Wide Band service, requirements of the concerned services become stricter, and networks of larger scales, complicated techniques and various devices appear. As a result, network failure appears more frequently and complicatedly. Reflected network behavior features, traffic dynamic characteristic research becomes hot when researchers doing a lot of work on information supervision, congestion control, bandwidth distribution and performance evaluation of the high speed network [1]. Because of network's nature of dynamic, noise and unsteady state, the traffic is bursting and unsteady, which brings trouble to the research on traffic features and traffic analysis model [2], while the research is fundamental to the network performance management, traffic predication, and network structure design.

II. TRAFFIC ANALYSIS MODEL

The advantage of the traffic analysis model is that major factors affecting the network are showed concisely and precisely in it while some relatively subordinate ones are ignored. SRD (short range dependency) traffic models have been in a significant academic position and many models exist, including Markov model and regression model. While the Markov model mainly includes interrupt Poisson procedure (IPP) model, ON/OFF model, ASRP (altered states refresh process) model, Markov-modulated Poisson process model (MMPP), and Markov-modulated fluids model (MMFM); the regression model includes auto-regressive (AR) model, discrete auto-regressive (DAR) model, auto-regressive moving average (ARMA) model, auto-regressive integrated moving average (ARIMA) model and transformation expansion sampling (TES) model. Although the models have their own characteristics, their structures are all exponential decaying. What more, the decaying speed is much faster than what have observed, which leads the difference between the result of queuing analysis and evaluation and the practical result. According to the new analysis on practical traffic, it is found that no matter in LAN or WAN, the traffic is with the nature of high variability and

bursting in a large time scale, while the high variability could be described by LRD (long-range dependency).

A. Method of Analysis on Self-similarity/Fractal

The simplest LRD traffic model is based on self-similar process model [3]. Its variance function decays by hyperbolic form. Self-similar process and asymptotically self-similar process act better than Poisson process does at describing the feature of practical traffic. The models based on self-similarity could reflect the features of LRD only by Hurst parameter. Because the traditional SRD models could not reflect self-similar traffic's affection to the network, LRD models complied with network businesses are required. LRD models are mainly with the following categories:

1) Fractional Autoregressive Integrated Moving Average (F-ARIMA) model

F-ARIMA (p, d, q) is the expansion of ARIMA (p, d, q). It is a compensation to ARIMA on the description ability. Their mathematical expression is as the following:

$$\phi(B)\nabla^d X_t = \theta(B)\varepsilon_t. \quad (1)$$

In the expression, ε_t is Gaussian white noise; $\phi(B)$ is the polynomial of mode p while $\theta(B)$ is the one of mode q . B is the delay operator in $\nabla = 1 - B$, which is $BX_k = X_{k-1}$. The operator $\nabla^d = (1-B)^d$ could be expanded as

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k B^k. \quad (2)$$

While

$$\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}. \quad (3)$$

$\Gamma(x)$ is gamma function.

$$\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p), \theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q). \quad (4)$$

The difference between them is that parameter d in ARIMA is positive integer, while the one in F-ARIMA is $0 < d < 0.5$.

Under extreme conditions, the self-similar function and spectral density function of FARIMA(p, d, q) process have the same structure as that of ARIMA($0, d, 0$), in other words, When it is $0 < d < 0.5$, FARIMA(p, d, q) is long range dependence.

When it is $0 < d < 0.5$, the self-similarity of $FARIMA(0, d, 0)$ is

$$r(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)} \sim \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}, (k \rightarrow \infty). \quad (5)$$

From the nature of self-similarity $r(k) \sim L(k)k^{-\beta}$ we know that the processes of $FARIMA(0, d, 0)$ and $FARIMA(p, d, q)$ can describe LRD of self-similarity of traffic at the premise of $H = d + 0.5$.

In building FARIMA model, by means of aggregated variance-time analysis, R/S analysis, spectrum analysis algorithm based on periodogram etc., we can get the estimate of H from the dynamic Hurst parameter and the estimate of d from $H = d + 0.5$ to achieve the pretreatment of dynamic data and the structural analysis of LRD of these data. By F-ARIMA, the functions of LRD and SRD features could be observed, which could not be realized by AR and ARIMA.

2) FBM

In order to research on the physical process of LRD, Mandelbrot expanded Brownian Motion (BM) to Fractional Brownian Motion (FBM). FBM could be realized from FBM by the following:

$$fB_t = \int_0^t (t-u)^{H-0.5} dB(u). \quad (6)$$

The self-similar function of FBM is

$$r(k) = \frac{1}{2} \left[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta} \right]. \quad (7)$$

When $k \rightarrow \infty$, $r(k) \sim k^{-\beta}$. So FBM is LRD and $H = 1 - \beta/2$.

FBM has its own advantages: It is strictly under mathematical theory. Its structure is simple, only uses 3 parameters of velocity, variance and Hurst. So it is usually used in theory analysis. However, FBM is strictly self-similar and good at modeling on LRD data but not modeling on SRD at the same time. Therefore, FBM could not describe the practical situation precisely. Its Gaussian nature leads an imprecise description on in-negative signal.

3) Superposition model in which the sojourn time complies with heavy-tailed distributed (HTD) ON/OFF sources(SM)

LSD sequence could be produced by SM. The theory is: every data source alternates between ON and OFF. When the data source transmits cell under ON with a velocity, if the sojourn time complies with a HTD stochastic variable, the variance is infinite; no cell is transmitted under OFF, and the sojourn time is IID (Independent identically distribution) stochastic variable. The traffic bursting length, grouping arrival time interval and connecting duration time is fractal, and the corresponding probability distribution is heavy-tailed. That is called Noah effect. Such data sources super posited together and formed a gradually self-similar aggregated traffic, while the self-similar parameter is $H = (3 - \beta)/2 > 0.5$. That is the Joseph effect in fractal theory. When the data sources is plenty enough, the arrival groupings in unit slot time equals with $M/G/\infty$ model.

The advantage of the model is that it simplified aggregation traffic to a LRD one, which could be used to describe LRD.

The model is usually used in network emulation and traffic aggregation. However, under HTD, it is difficult to be dealt with mathematically and it is not clear whether it could be used to describe SRD.

B. Multi Fractal

With the scale analysis on the traffic of LAN and WAN, Feldmann group found that it is not enough to use only a Hurst parameter to describe the features of network traffic [4]. It is found that the gradual self-similarity in a large time scale, when observing and analyzing WAN traffic, could be described precisely by fractal/self-similarity; while the partly special characters in small-time scale (less than hundreds million second) could not, but only by multiracial. Because the partly special characters in small time scale of the traffic closely relate to the operation and control of the network, the research on multi fractal could reflect the interrelationship when different network connection happens in different protocols.

The multi fractal feature of WAN aggregation traffic was firstly discovered by Riedi and Vehel [5]. Then Feldmann, Gilbert and Willinger held further research on it and found it was probably caused by the conservative cascade operation system of IP network [4].

By the method of wavelet analysis, the partly special characters of WAN traffic and Holder index number could be described by wavelet parameter. The method of wavelet analysis is a technique of scale detecting, marketing and utilizing. It is good at analyzing unsteady signals, both in general and in details. DWTM (DWT-based wavelet-coefficients model), based on DWT (discrete wavelet transformation), is typical to modeling by wavelet. Its nuclear theory is to aggregate wavelet coefficient and scale coefficient by analyzing wavelet coefficient in the traffic and contradictorily wavelet transferring.

1) DWTM on DWT

Wavelet analysis principal:

Any real number of $(a, b; a, b \in \mathbb{R}, a > 0)$ calls the following function

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (8)$$

as a continuous wavelet function which produced by wavelet generating function $\psi(t)$ and dependent on parameter (a, b) , short for wavelet. Obviously $\int_{\mathbb{R}} \psi_{a,b}(t) dt = 0$. So according to any function or signal $f(t)$, it changes to:

$$W_f(a, b) = \int_{\mathbb{R}} f(t) \overline{\psi_{a,b}}(t) dt = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(t) \overline{\psi}\left(\frac{t-b}{a}\right) dt. \quad (9)$$

In the expression, a is elastic factor and b is translation factor. Wavelet transformation is an invariant-preserving transformation. There is no loss of power in the transformation. The elastic factor a determines the scale of signal observation while translation b is used to help choose the position on which the signal is analyzed [6]. The square of the wavelet range index is a scalegram, with a description of scale a 's signal energy value around time b . While the energy value of the signal in scale a , also named wavelet energy spectrum, is as the following:

$$E(a) = \int_R |W_f(a, b)|^2 db. \quad (10)$$

The wavelet energy spectrum of signal $f(t)$ is weighted average from its power spectrum. The weight function is the wavelet function's power spectrum:

$$E(a) = \frac{1}{2\pi} \int_R S_f(\omega) S_\psi(\omega) d\omega. \quad (11)$$

The wavelet power spectrum provides a method to describe the whole process of signal. Its inverse transformation is

$$f(t) = \frac{1}{c} \iint_{R^2} \frac{1}{a^2} W_f(a, b) \psi\left(\frac{t-b}{a}\right) da db. \quad (12)$$

While

$$C = \int_R |\overline{\psi(\omega)}|^2 |\omega|^{-1} d\omega < +\infty. \quad (13)$$

Because the continuous wavelet transformation is a redundancy transformation, the correlation of wavelet between the two spatial points adds up to the difficulty of analysis. In applying the wavelet transformation to the network traffic analysis, dispersive wavelet transformation is demanded so that the dispersive data signals can be better processed. Dispersive wavelet transformation is to disperse the extension factor a and shift factor b of the continuous wavelet transformation.

When

$$a = a_0^j, b = kb_0 a_0^j, a_0 > 1, b \in R. \quad (14)$$

k and j are integral number, dispersive wavelet $f(t)$ changes to:

$$W_f(j, k) = a_0^{-j/2} \int_R \overline{\psi(a_0^{-j} t - kb_0)} f(t) dt. \quad (15)$$

When $a_0=2$, $b_0=1$, $W_f(j, k)$ transforms to discrete dyadic wavelet transform.

That is

$$W_f(j, k) = 2^{-j/2} \int_R \overline{\psi(2^{-j} t - k)} f(t) dt. \quad (16)$$

The multi scale analysis of signals could be realized by the change of j .

The basic algorithm of DWWM is to Haar DWT factoring the original traffic data to get the related elastic factor $\{a'(k)\}$ and wavelet coefficient $\{d'(k)\}$. And then record the scale factor and wavelet coefficient in the large scale, analyze the distribution features of the wavelet coefficient and record the parameters. It could be found that the wavelet coefficient is Gaussian normal distributed around average zero. And then to describe the distribution feature by standard deviation $std(j)$ and the related bound.

MWM is a model of multiplication. Its aggregated signal could be performed by formula of multiplication. Because its

data is non-negative and LRD with the limitation, and the model itself is not Gaussian, it could describe the practical traffic precisely. While for the modeling of LRD traffic, only 3 parameters needed and its bound is gradual logarithmic normal distributed. It could describe the LRD and SRD in the traffic. It could be used easily in practice for its simple calculation [5].

2) Cascade Model

Cascade process is fundamental in the field of multi fractal. It segments a target to minor fragment by some geometric regulation, and distributes the measurement to those fragments by another certain regulation. Internet's working process is similar to a conserve cascade process (if the extra protocol existed in the layers ignored). So the complicated scale features of practical traffic could be reflected by the traffic model on cascade process. For example, a Web conversation would be segmented to many HTTP requirements; every HTTP requirements could produce many TCP links; each TCP links have many IP flows; and every IP flow is consisted with data packets. This process could be regarded as a random nearly conservation cascades process.

The basic theory of cascade process modeling is: To divide the initial flow data sequence into odd-even subsequence; calculated for the higher sequence and the related partition coefficient sequence; repeatedly segment and aggregate the new sequence and keep calculating the related distribution coefficient sequence until only one number is left. That number is the sum of all factors in the beginning sequence. By the normal distributed average and standard deviation, the random distribution sequence to calculating the sequence of the average, standard deviation, maximum, minimum and the maximum dimension existed. The features of self-similarity and multi-fractal could be known by calculating structure function, partition function and linear feature detection.

III. CONCLUSION

Large scale network is a complicated non-linear system. Nowadays the research on traffic behavior is undergoing and many problems are waiting to be solved. In order to understand the facts of network's microcosmic phenomenon, we need to know the network and its regulation: we need to search for the internal relations between modeling analysis and macroscopic flow model when we add new protocol features or new flow measurements to the realization of the protocols [7]. In the future, the traffic analysis will take the multi features of the flow into consideration and could describe the development of any flow more precisely.

ACKNOWLEDGMENT

In the process of bringing this thesis into its present form, the authors have been fortunate to benefit much from the projects funded as follow:

- 1) Project of the Scientific and Technological Study of Chong Qing Municipal Education Commission(KJ081116).
- 2) Project of the Scientific and Technological Study of Chong Qing Three Gorges University(2007-sxxyn1-10).

REFERENCES

- [1] Nevil Brownlee, Claffy Kc. Internet Measurement [C] //Proc of IEEE Internet Computer. Piscataway, NJ : IEEE, 2004 :30-33.

- [2] Xie Gaogang , Zhang Guangxing , Yang Jianhua , et al. The survey on traffic of metro area network with measurement on-line [C] //Proc of the 20th Int Teletraffic Congress.Berlin: Springer,2007: 666-677.
- [3] Willinger W, Paxson V. Where Mathematics Meets The Internet. In Notices of the American Mathematical Society, Sept.1998,Vol.45,No.8:pp.961-970.
- [4] Feldmann A, Gilbert A C, Willinger W, Kurtz T G.The Changing Nature of Network Traffic: Scaling Phenomena. Computer Communication Review 28, April 1998, No.2.
- [5] Riedi R H., Crouse M S, Ribeiro V J, Baraniuk R G. A multifractal wavelet model with application to network Traffic [J]. IEEE Transactions on Information Theory, April 1999 45 (3): 992 – 1018.
- [6] Qiao Yi, Skicewicz Jason , Dinda Peter. An empirical study of multiscale predictability of network traffic [C] //Proc of the 13th IEEE Int Symp on High Performance Distributed Computing. Piscataway, NJ : IEEE , 2004 : 66-76.
- [7] Subhabrata Sen, Jia Wang. Analyzing peer-to-peer traffic across large networks [J] . IEEE/ACM Trans on Networking, 2004, 12 (2) : 219-232.