Heterogeneous Wireless Network Traffic Load Estimation based on Chaos Theory

Xue Han[1,2,3], Yucheng Zhang[1,2], Liang Huang[1,2,3], Jinglin Shi[1,2]
1. Wireless Communication Technology Research Center, Institute of Computing Technology, Chinese Academy of Sciences
2. Beijing Key Laboratory of Mobile Computing and Pervasive Device
3. Graduate University of Chinese Academy of Sciences, Beijing, China
Email:{hanxue, zhangyucheng, huangliang, sjl}@ict.ac.cn

Abstract—**The co-existence of multiple wireless networks has made the wireless network environment complex and heterogeneous. Traffic flow characteristics study and effective traffic prediction algorithm selection are very important for optimal allocation of network resources, network protocol design and improvement of service quality in this heterogeneous wireless network. This paper uses non-statistical method which is based on chaos theory to make analysis of heterogeneous wireless network traffic flow characteristics. On that basis, a traffic model which uses phase space reconstruction algorithm on multivariate traffic time series is proposed and used for traffic prediction. Simulation results show that the proposed traffic prediction model can achieve better prediction performance.**

I. INTRODUCTION

The continuous emerging of wireless broadband communication technologies make wireless communication network of the future more and more heterogeneity [1][2]. The number of network user increases exponentially following with massive online network services. This will pose challenge to heterogeneous wireless network's transmission capacity and its ability to ensure QOS of services. Accurate prediction of network traffic is one solution to this challenge. Network traffic model is a basic research in the field of traditional network. It is an important guidance for optimal allocation of network resources, congestion control, load balancing and etc. In order to realize cooperation among heterogeneous networks, it is necessary to do research on how to build the traffic model and on the method of traffic prediction.

In terms of traffic model research, Poisson model is recognized as the traditional circuit-switched network traffic model. In Internet and wireless network, as the traffic properties change significantly, there is no such well accepted model.

Traffic analysis mainly base on Markov model, self-similar theory and statistical method. The author in [3] makes classification of the packets transferred in network via packet-level Hidden Markov Model. [4] uses Semi-Markov process to represent the behavior of mobile users and the resultant traffic pattern. The model proposed in that paper is used to make estimation of the resources in network and ensure quality-of-service to mobile users. By customizing the batch Markov arrival process (BMAP), [5] presents a synthetic traffic model for UMTS based on measured trace data. [6] is the first to use self-similar theory to reveal the scale characteristics of network traffic. It sets the traffic studies apart from the influence of poisson model and make the research enters a new historical stage. [7] presents a model of LAN traffic based on self-similar "on-off" function. [8] uses seasonal ARIMA models to forecast traffic workload which is important in planning, design, control and management of wireless network. Other studies that focus on the traffic flow characteristics are very helpful for traffic model establishment. For example, [9] considers network traffic as alpha/beta model in opposition to the aggregate modeling ideal. [10] uses NeTraMet to measure the size and lifetime of Internet streams. This method is used to characterize traffic distribution.

All the studies mentioned above focus on single network. They can't reflect the characteristics of heterogeneous wireless network. They are also not aimed at the traffic prediction properties and hence not suit for traffic forecast. In this paper, we adopt chaos theory based non-linear method to make analysis of heterogeneous wireless network traffic flow characteristics and propose a traffic model which uses phase space reconstruction algorithm on multivariate traffic flow time series for traffic prediction. This paper is organized as follows: In section II, the Traffic Model of Heterogeneous wireless network and the traffic prediction algorithm are introduced. In section III, a simulation scenario is set up to test the prediction performance of the proposed traffic model. The prediction result is compared with ARIMA prediction. Finally, a conclusion is drawn.

II. TRAFFIC MODEL DESCRIPTION

A. *Chaotic time series phase space reconstruction theory*

Chaos theory is an important branch of non-linear science. It reveals the unity of order and disorder in non-linear dynamical systems. Chaos time series is a univariate or multivariate time series obtained from a chaotic dynamical system. It seems to be stochastic from a micro point of view, but deterministic in macro viewpoint. The analysis of chaos time series is mainly to determine its strange attractor.

As one of the characteristics of chaotic system, the existence of chaotic strange attractor means that the chaotic system will fall into a deterministic trajectory. The idea of reconstruction of chaotic time series comes from the embedding theorem raised by [11]. The main point of Takens theory is that the phase space of a strange attractor can be reconstructed from a time series of data. Hence, the

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

embedding order of the chaos time series is estimated. As time goes by, two adjacent chaotic orbits will depart rapidly. But, as long as the evolution has not gone far, the two orbits are still not far away from each other. That's why the Takens theory can be used for precision short term forecast.

For a certain point in phase space, in order to forecast its next step, we have to find its nearest neighbors first. According to the next step of each nearest neighbor, the prediction can be made. That's the process of chaotic prediction. Consider a time series

$$\{x_q:\ q=1,2,...,N\} \quad (1)$$

where q denotes the number of samples. The observation time interval is constant. Using time-delay parameter τ and embedding dimension m to embed the time series into a m-dimension Euclidean space R^m, a set of embedded phase space vectors is as following:

$$X_q = (x_q, x_{q+\tau}, ..., x_{q+(m-1)\tau}) \quad (2)$$

The restored phase space trajectory is:

$$X = (X_1, X_2, ..., X_M)^T \quad (3)$$

Where $M = N-(m-1)\tau$.

The selection of τ and m has an impact on the precision of phase space restore. The method for determine them will be introduced in section II-B.

B. Heterogeneous wireless network traffic model

Consider the chaotic time series as equation (1) with τ and m as mentioned before, a m-dimension state space can be expressed as equation (2). The mapping from q to $q+1$ step has two equivalent expressions as following:

$$X_{q+1} = F(X_q), \quad \mathrm{F}: R^m \to R^m \quad (4)$$
$$x_{q+\tau} = f(X_q), \quad \mathrm{f}: R^m \to R$$

The relationship between vector function F and numerical function f is $F(X_q) = [x_q, x_{q+\tau}, ..., x_{q+(m-2)\tau}, f(X_q)]^T$. For chaotic system, function f is non-linear. The deterministic model of a given chaotic time series is to construct a function $\hat{f}$ as an approximation of f. This deterministic model can be used to make prediction. So, chaotic time series modeling and forecasting have the following steps:

Step1: Determine if a given time series as equation (1) is a chaotic one by calculating its maximum Lyapunov index.

Step2: Calculate τ and m if the time series is chaotic.

Step3: Build the deterministic model by extending one dimensional time series to higher dimension and reconstructing the phase space.

Step4: Adopt certain algorithm to make prediction based on the model established in Step3.

The method of calculating the maximum Lyapunov index mentioned in Step1 is as following:

Step1: Reconstruct phase space on base of τ and m.

Step2: Starting from a central point X_1, calculate the nearest neighbors of each phase space point X_q with limited short-term separation $d_q(0) = \min_j \| X_q - X_j \|, |q-j| > T$, where $\|.\|$ is Euclidean distance and T is mean period that can be got from Fourier transform. Calculate the distance $d_q(k)$ between X_{q+k} and X_{j+k} which denotes the pair of neighbors in phase space after k discrete time steps. That means $d_q(k) = \| X_{q+k} - X_{j+k} \|, k = 1,2,...,\min(M-q, M-j)$, $M = N-(m-1)\tau$ is the number of phase points after phase space reconstruction.

Step3: For each q, calculate $y(k) = \frac{1}{M}\sum_{q=1}^{M} \ln(d_q(k))$. $d_q(k)$ is not zero and M is the number of such $d_q(k)$ for a certain k.

Step4: Draw $y(k)$ curve with k. Use the least square method to calculate the slope of the curve which is the maximum lyapunov index λ. If λ>0, then the time series is chaotic.

C-C method[12] is adopted for τ and m calculating. C-C method is as following:

Step1: Calculate the standard deviation σ for the given time series.

Step2: Calculate $\Delta\bar{S}(t)$ and $S_{cor}(t)$ according to formulas below:

$$\begin{aligned} \bar{S}(t) &= \frac{1}{16}\sum_{m=2}^{5}\sum_{j=1}^{4} S(m, r_j, t) \\ \Delta\bar{S}(t) &= \frac{1}{4}\sum_{m=2}^{5} \Delta S(m,t) \\ S_{cor}(t) &= \Delta\bar{S}(t) + |\bar{S}(t)| \end{aligned} \quad (5)$$

Where $r_j = \frac{j\sigma}{2}$ and t is a natural number less than or equal to 200. $S(m,r,t)$ and $\Delta S(m,t)$ can be got as below:

$$S(m,r,t) = \frac{1}{t}\sum_{s=1}^{t}\left[C_s(m, r_j, t) - C_s^m(1, r_j, t)\right] \quad (6)$$

$$\Delta S(m,t) = \max\{S(m, r_j, t)\} - \min\{S(m, r_j, t)\} \quad (7)$$

Step3: Draw the curves of $\Delta\bar{S}(t)$ and S_{cor} The corresponding value of t when $\Delta\bar{S}(t)$ first reaches minimum is delay-time parameter τ .The corresponding value of t when S_{cor} is minimum is $\tau_w = (m-1)\tau$ .So, we can get τ and m .

Until now, we have only considered phase space reconstruction using single variable time series. In heterogeneous wireless network environment, traffic from sub-networks has different contribution to the total network traffic. Factors that have effect on sub-network's contribution to heterogeneous network include channel quality, the place where traffic flow time series is observed, time granularity used for traffic flow measurement, delivery method (ATM,FR,TDM,IP) and etc. Therefore, phase space reconstruction based on single variable time series can't describe the evolutionary trajectory of heterogeneous wireless network accurately and hence do harm to the precision of traffic prediction. In fact, single variable time series is a special case of multivariate time series. So, it is easy to apply the phase space reconstruction theory to multivariate time series.

Consider a heterogeneous wireless network which includes P sub-networks. The total traffic is the sum of these P sub-network traffic. So, the time series extend to P variables:

$$\{x_q\}_{q=1}^{N} = \{(x_{1,q}, x_{2,q}, ..., x_{P,q})\}_{q=1}^{N} \quad (8)$$

It is a measurement of P continuous variables $x(t) = \{x_1(t), x_2(t), ..., x_P(t)\}$. If each time series is a chaotic one, we can get the heterogeneous wireless network traffic model according to the multi-dimensional phase space reconstruction theory in [13]:

$$\begin{bmatrix} S_{q_0} = \{x_{1,q_0}, x_{1,q_0+\tau_1}, ..., x_{1,q_0+(m_1-1)\tau_1}, ..., x_{P,q_0}, x_{P,q_0+\tau_P}, ..., x_{P,q_0+(m_P-1)\tau_P}\} \\ ... \\ S_{q_i} = \{x_{1,q_i}, x_{1,q_i+\tau_1}, ..., x_{1,q_i+(m_1-1)\tau_1}, ..., x_{P,q_i}, x_{P,q_i+\tau_P}, ..., x_{P,q_i+(m_P-1)\tau_P}\} \\ ... \\ S_N = \{x_{1,N}, x_{1,N+\tau_1}, ..., x_{1,N+(m_1-1)\tau_1}, ..., x_{P,N}, x_{P,N+\tau_P}, ..., x_{P,N+(m_P-1)\tau_P}\} \end{bmatrix} \quad (9)$$

In the above matrix, $q_0 = \max_{1\le p\le P}(m_p - 1)\tau_p + 1$, τ_p and m_p are the delay-time parameter and embedding dimension of the p_{th} time series. $p = 1, 2, ..., P$ (If p =1, then it is a single variable time series). The embedding dimension of this multi-variate deterministic model is $m = m_1 + m_2 + ... + m_P$.

C. Heterogeneous wireless network traffic local prediction algorithm

The basic idea of local prediction method is that choose several phase points from the neighborhood of a point at certain time. Then get fitting function f on base of these neighborhood phase points. Function f can be either linear or non-linear. In this paper, we adopt widely used weighted zero-rank local prediction algorithm to make heterogeneous wireless network traffic forecast. This algorithm for single variable time series is introduced first and will be extended to multivariate later.

For a given time series as equation (1) with τ and m, its phase space points is as following:

$$\begin{aligned} x_1 &= (x_1, x_{1+\tau}, ..., x_{1+(m-1)\tau}) \\ x_2 &= (x_2, x_{2+\tau}, ..., x_{2+(m-1)\tau}) \\ &... \\ x_N &= (x_N, x_{N+\tau}, ..., x_{N+(m-1)\tau}) \end{aligned} \quad (10)$$

If we can get the next phase point of x_N which is denoted as $x_{N+1} = (x_{N+1}, x_{N+1+\tau}, ..., x_{n+1})$, then the prediction value is x_{n+1} in the next time interval. That is because x_{N+1} is the evolution result of x_N. To get x_{N+1}, we have to find the nearest neighbors of x_N as $(x_N)_i, i = 1, 2, ...L$. The average evolutionary behavior of these L number of phase space points can express the evolutionary behavior of x_{N+1}. The estimation of x_{N+1} is $\hat{x}_{N+1} = \frac{1}{L}\sum_{i=1}^{L}(x_N)_i$. The last component of $\hat{x}_{N+1}$ is the prediction of x_{n+1}

Let's further consider the effect of the distance between central point and each nearest neighborhood point on prediction result. In fact, each nearest neighborhood point affects the central phase point differently. The weighed zero-rand local prediction algorithm seems the distances mentioned above as a fitting parameter of evolution equation. As a result, the prediction accuracy can be improved. The modified prediction model is as following:

$$\hat{x}_{N+1} = \frac{\sum_{i=1}^{L}(x_N)_i e^{-A(d_i - d_{\min})}}{\sum_{i=1}^{L} e^{-A(d_i - d_{\min})}} \quad (11)$$

Where d_i is the Euclidean distance between the nearest neighbor point and central point. $d_{\min}$ is the minimum value of d_i. L is the number of nearest neighbor. A is a parameter that can be revised and is generally taken as $A \geqslant 1$. It can be seen from this model that the closer the neighbor point is away from the central point the greater it is weighted in prediction. Now we give the weighed zero-rand local prediction algorithm which is suitable for the heterogeneous wireless model mentioned in chapter II-B for network traffic prediction. Suppose S_{q_i} is the central point, its next step is $S_{q_{i+1}}$.

First of all, we have to get the L neighbor points $S^{\min}_{q_i,j}$, j=1,2,…, L which have nearest Euclidean distance from S_{q_i}. Each Euclidean distance is denoted as d_j, j=1,2,…, L. Let $d_{\min}$ be the minimum value of d_j, then the weight of each $S^{\min}_{q_i,j}$ is as following:

$$P_j = \frac{e^{-A(d_j - d_{\min})}}{\sum_{j=1}^{L} e^{-A(d_j - d_{\min})}} \quad (12)$$

Finally, since $S_{q_{i+1}} = \{(x_{1,q_{i+1}}, x_{1,q_{i+1}+\tau_1}, \ldots, x_{1,q_{i+1}+(m_1-1)\tau_1}), \ldots,$ $(x_{P,q_{i+1}}, x_{P,q_{i+1}+\tau_P}, \ldots, x_{P,q_{i+1}+(m_P-1)\tau_P})\}$, we can get the multi-variate traffic prediction $\hat{S}_{q_{i+1}} = \{\hat{x}_{1,q_{i+1}}, \hat{x}_{2,q_{i+1}} \ldots, \hat{x}_{P,q_{i+1}}\}$. So the heterogeneous wireless network traffic prediction is $\hat{S}_{i+1} = \sum_{p=1}^{P} \hat{x}_{p,q_{i+1}}$.

III. SIMULATION

In this section, a simulation scenario is set up to test the prediction performance of the traffic model proposed in this paper. By triggering heterogeneous wireless network services, the network traffic time series can be got. This simulation scenario is a 200m*400m rectangular area with UMTS network and a 802.11b WLAN. First we introduce how the traffic is obtained. Then validate if it is chaotic. Finally we'll make prediction on base of the algorithm mentioned in section II-C. The prediction result is compared with ARIMA prediction.

A. Obtain the heterogeneous wireless network traffic

In this simulation scenario, properties such as charges, power and bandwidth are static. The arrival interval of services is supposed to obey Poisson distribution. Services trigged by UMTS network include voice, E-mail, web and short message. WLAN services include ftp, web, fax, video. The detailed parameters are listed in table I.

TABLE I. HETEROGENEOUS WIRELESS NETWORK SIMULATION PARAMETERS

Sub-network	Service	Arrival rate （10^{-5}）	distribution	Average duration （s）
UMTS	voice	20	Poisson	180
	E-mail	5	Fixed length	10
	Web	1	Poisson	60
	Short message	20	Fixed length	5
WLAN	ftp	20	Poisson	20
	Web	20	Poisson	1800
	fax	1	Poisson	100
	video	5	Poisson	1800

Suppose that there exists a router through which all the traffic of that heterogeneous wireless network goes out to other networks. Then the total traffic can be obtained from that router. UMTS and WLAN traffic time series can be got separately. These time series are normalized according to the following function:

$$l'(n) = \frac{l(n) - \frac{1}{N}\sum_{n=1}^{N} l(n)}{\max[l(n)] - \min[l(n)]} \quad n = 1,2,\ldots,N \qquad (13)$$

The normalized network traffic time series are described in Fig1.

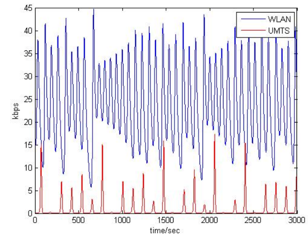

Fig 1. Traffic load time series

B. Time series Chaotic analysis

Use the C-C method described in section II-B to calculate τ and m of UMTS and WLAN time series. The curves of $\Delta\bar{S}(t)$ and S_{cor} are as Fig2.

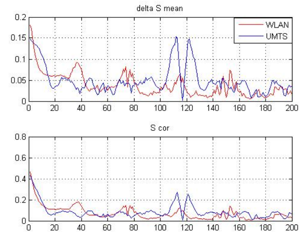

Fig 2. Curve of delta S mean and S cor

According to Fig2, it can be found that for UMTS, τ is 18 and m is 7. For WLAN, τ is 45 and m is 4.

Now the maximum lyapunov indices is calculated using method described in section II-B. The least square regression curves are as Fig3 and Fig4.

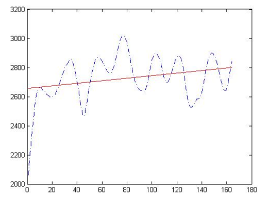
Fig 3. WLAN time series lyapunov index curve

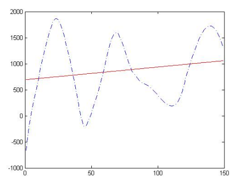
Fig 4. UMTS time series lyapunov index curve

From Fig3 and Fig4 It can be seen that the maximum lyapunov indexes for both time series are greater than zero. So the heterogeneous wireless network is chaotic. Table II sums up the parameters calculated in this section.

TABLE II. LIST OF PARAMETERS

Time	Embedding	Delay-time	maximum

series type	dimension m	parameter τ	lyapunov index
WLAN	4	45	2.4474
UMTS	7	18	0.8771

C. Prediction and results analysis

Apply the zero-rank local prediction algorithm introduced in section II-C to the traffic time series for prediction. The last phase space point of time series is chosen as the central point. The number of nearest neighbor points is limited to 10. According to section III-B, τ and m for UMTS is 18 and 7, τ and m for WLAN is 45 and 4. Compared the prediction result with that getting from the ARIMA prediction model proposed in [8]. The prediction results from both models and the actual traffic are show in Fig5. Prediction deviation which is calculated according to

$$err = \frac{l_{\exp}(t) - l_{real}(t)}{l_{real}(t)} \qquad (14)$$

is shown in Fig6. According to Fig6, The prediction deviation of the method proposed in this paper is about 25% of ARIMA.

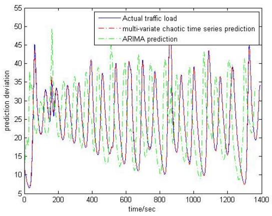

Fig 5. Traffic prediction results

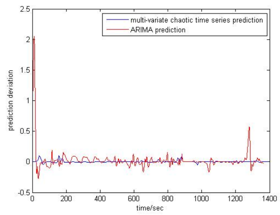

Fig 6. Prediction deviation comparison

IV. CONCLUSTION

Nowadays, people's demands for wireless services with high quality at anywhere and anytime can't be satisfied by one single network. Therefore, the concept of heterogeneous wireless network integration comes into being. In this paper, we make analysis of the network traffic properties by taking the various sub-networks in heterogeneous wireless network as a variable in multivariate phase space reconstruction theory. We first prove that the heterogeneous wireless network traffic is chaotic. On that basis, we propose a traffic prediction model. The simulation results show that our traffic model have good prediction performance.

V. ACKNOWLEDGEMENT

This work was funded by Beijing Natural Science Foundation Major Project (4110001), Major Project of the Ministry of Science and Technology of China (2012ZX03003005-04) and Co-building Program of Beijing Municipal Education Commission.

REFERENCES

[1] Chen Yiping ,Yang Yuhang "A New 4G Architecture Providing Multimode Terminals Always Best Connected Services", IEEE Wieless Communication 2007 April.

[2] Y. Yuan, Z. Zhang, J. Li, J. Shi, J. Zhou, G. Fang, and E. Dutkiewicz,Extension of SCTP for concurrent multi-path transfer with parallel subflows, in Proc. IEEE Wireless Communications and Networking Conference, pp. 1C6, Apr. 2010.

[3] Alberto Dainotti, Walter de Donato, Antonio Pescap,Pierluigi Salvo Rossi, "classification of Network Traffic via Packet-Level Hidden Markov Models", 2008 GlobalCom.

[4] Hisashi Kobayashi, Shun-Zheng Yu, Brian L. Mark, "An integrated mobility and traffic model for resource allocation in wireless networks", WOWMOM '00 Proceedings of the 3rd ACM international workshop on Wireless mobile multimedia 2000.

[5] Klemm, A., Lindemann, C., Lohmann, M., "Traffic modeling and characterization for UMTS networksGlobal Telecommunications Conference", 2001. GLOBECOM '01. IEEE.

[6] W.E Leland et al. "On the self-similarity Nature of Ethernet Traffic", IEEE Trans. Networking, Vol.2,No.1,pp.1-15,Feb.1994.

[7] Janusz Kolbusz, Stanis'aw, Paszczyn'ski, Senior Member, IEEE, and Bogdan M. Wilamowski, Fellow, IEEE, "Network Traffic Model for Industrial Environment", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, Vol. 2, No. 4, Nov. 2006

[8] Yantai Shu, Minfang Yu, Jiakun Liu, Yang, O.W.W., "wireless traffic modeling and prediction using seasonal ARIMA Models", Communications, 2003. ICC '03.

[9] Sarvotham, S., Riedi, R., and Baraniuk, R., "Connection-level analysis and modeling of network traffic", In Proc. 1st ACM SIGCOMM Workshop on Internet Measurement .San Francisco, California, USA, November 01 - 02, 2001.

[10] Brownlee, N. and Claffy, K.C., "Understanding Internet traffic streams: dragonflies and tortoises", Communications Magazine, IEEE, vol.40, no.10, October 2002, pp. 110-117.

[11] Takens F. "Determining strange attractor in turbulence"Lecture notes in Math,1981,898361-381

[12] DHeyman, T .VLakshman, ATabatabaiet a1. ,"Modeling Teleconference Traffic from VBR Video Coders", IEEE International Conference on RecordServing Humanity Through Communications, V013, May 1994, PP1744-1 748.

[13] Liangyue Cao, Alistair Mees, Kevin Judd, "Dynamics from multivariate time series". Physica D, 1988,121(1-2):75-88.