The Nonlinear Mechanism of Phase Transition in Computer Networks*

Li Yi-Peng[1], Huang Yi-Hua[2], Wang Lei[1], and Ren Yong[1]

[1] Department of Electronic Engineering, Tsinghua University, P.R. China
yp-li05@mails.tsinghua.edu.cn
[2] Department of Electronic and Communication Engineering, SUN YAT-SEN University, P.R. China

Abstract. In this paper, the nonlinear mechanism of phase transition in computer networks is analyzed, and a distributed proxy approach is introduced to improve network performance based on the two-dimensional coupling model. Theoretical analysis figures out that the nonlinear mechanism of router is the essential reason of network performance phase transition. Simulation results reveal that the extreme clustering characteristic of web access behavior gives arise to left-shift of phase transition critical compared with regular networks; after distributed proxy approach is employed, right-shift of the phase transition critical illustrates performance improvement. Finally, several important issues are mentioned.

Keywords: nonlinear mechanism, phase transition, distributed proxy.

1 Introduction

The rapid development of Internet applications brings more convenience for users collecting information, such as web network and typical P2P file sharing systems, which have become the majority traffic contributors to infrastructure. Millions of computers around the world attach to the Internet through many autonomous regional networks of routers, which interconnect through backbone networks of routers in a distributed, hierarchical fashion. All of these make Internet topology more complex and have negative impacts on infrastructure performance, e.g. poor transmission efficiency and network congestion.

To analyze and solve these problems, relative concepts and theories in statistic physics have already been introduced to characterize the collective dynamics of Internet traffic. Phase transition, for example, has been used to characterize the internet performance fluctuation, distinguishes the network performance status as free-flow and congestion. In the previous relevant research, Willinger et al. [1] provided a simple

* This work is supported in part by the National Nature Science Foundation of China (NSFC) grants 60672142, 60772053, the National Basic Research Program of China (973 Program) grants 2007CB307100 / 2007CB307105, and by the China Postdoctoral Science Foundation grants 20080430400.

J. Zhou (Ed.): Complex 2009, Part I, LNICST 4, pp. 560 – 568, 2009.

physical explanation for self-organized criticality of Internet traffic, caused by multiple ON/OFF processes. Then Csabai [2] and Takayasu *et al.* [3] characterized Internet traffic statistics property according to spectral characteristics, shown by ping sequences. In 2007, Hu and Wang [4] have extended phases of network performance, they found the fundamental diagram of flow against density, hysteresis inside, and classified the traffic flow with four states: free flow, saturated flow, bistable and jammed. All this kind of researches supply valuable theoretical basis and simulation supports, which is likely to contribute to better network engineering and management.

Internet, as a typical huge complex system, different applications running on it and the corresponding user behavior has shown complicated statistic property. Generally, the resulting performance fluctuation can be dependent on two important factors. First, the existing nonlinear relationship between router input and output, router interactions due to protocol, make Internet infrastructure being dynamic status in processing millions of packets. The critical status, corresponding to the phase transition point, indicates the decline in network performance. Second, clustering characteristic of user access behaviors, shown by traffic analysis, aggravates the nonlinear relationship mentioned and interactions between routers, leading to deterioration of network performance. Unfortunately, the nonlinear mechanism, the essential reason of phase transition, has never been discussed [2-5] to our best knowledge. Hence, it is important to study the nonlinear dynamic mechanism of network phase transition, and introduce a relative improving approach.

In this paper, a novel two-dimensional coupling model is proposed to describe the huge and complex link relations in computer network. Then we provide theoretical analysis to the nonlinear mechanism of phase transition and the distributed proxy approach. Based on the novel model, simulations reveal the negative effect of the nonlinear mechanism and the network performance improvement brought by employing the distributed proxy.

The rest of the paper is organized as follows. In Section 2, we describe the novel two-dimensional couple model in detail. In Section 3, we give a theoretical analysis to the nonlinear mechanism and the distributed proxy approach, which is also introduced into the two-dimensional coupling model. Section 4 shows extensive simulations by using the modified model in two different network sizes. We conclude in Section 5.

2 Two-Dimensional Coupling Model

In 1999, two-dimensional cellular automation [6] is used in network modeling for the first time, which became a useful tool to characterize the internet collective dynamics behavior [7]. For one thing, its non-periodic boundary condition makes the spatial distribution of packets unbalanced, which means some central nodes more congested than the others. For another, this kind of model is always with the default configuration that nodes are homogenous [8], which can not describe the distributed access behavior of users accurately. To describe the huge and complex link relations in computer networks, we proposed a novel two-dimensional coupling model (see Fig.1).

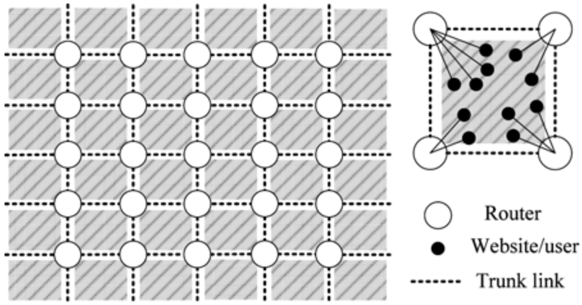

Fig. 1. Two-dimensional cellular automation model

In the two-dimensional coupling model, different websites or source providers and intensive users locate in shadow grid, connecting to the router nodes randomly. Routers, which follow the periodic boundary conditions [9], build up the backbone network, and route packets according to the shortest path routing strategy [10]. *L* is the size of the backbone network; means there are *L* routers in the horizontal and vertical direction of the backbone network.

The position of each router is denoted by a discrete space variable $\hat{r}$:

$$\hat{r} = i\hat{c}_x + j\hat{c}_y \tag{1}$$

Let K_r be the number of websites or users connecting to router $\hat{r}$, then the corresponding website or user position can be denoted as follow:

$$\hat{r}_k = i\hat{c}_x + j\hat{c}_y + k\hat{c}_z \tag{2}$$

where $\hat{c}_x$, $\hat{c}_y$, $\hat{c}_z$ are Cartesian unit vectors, and $i, j = 1,\cdots,L$, $k = 1,\cdots,K_r$.

The dynamics behavior of users in the model is governed by the parallel update with discrete time step. During the progress of model evolution, users create new packets with zero life time, which are forwarded by routers to some selected destination. All routers work as FIFO (First Input First Output). During each time step, the processing of packet transmission is shown as follows.

(1) Each user node sends one packet with probability *p* independently;
(2) According to FIFO rule, each router routes one packet in its buffer to destination by one hop, the next hop router queue length increases by 1, the routed packet lifetime increases by 1;
(3) If the packet is routed to the destination, the relative record ends.

At the time step k, if queue length of router $\hat{r}$ is $q(\hat{r},k)$, the sum queue length of all routers is

$$Q(k) = \sum_i q(\hat{r}_i, k)\ ,\ i = 1, \cdots, L^2 \tag{3}$$

The router's buffer size is set to

$$B_{Router} = 2 \times \sum_i K_{\hat{r}_i}\ ,\ i = 1, \cdots, L^2 \tag{4}$$

In the case of buffer overflowing, routers will drop packets until there is free space in the buffer.

3 The Nonlinear Mechanism of Phase Transition and Improvement

Research in [7] shows that the phase transition of network performance is due to the nonlinear interaction of routers. And our preliminary studies prove that extreme clustering characteristic of web access behaviors gives arise to left-shift of phase transition point [11]. But all these studies are still limited to large-scale or medium-scale.

Consider a simple example as Fig.2 shows.

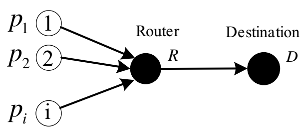

Fig. 2. Simple packet forwarding system

In this simple system, source nodes $1, \cdots, i$ send packets to destination D with independent probability p_i, which will increase from 0 to 1 gradually. Router R, queue length noted by Q_R, can output one packet per time step.

At the time step k, the expectations input and output of R in the sense of probability are:

$$E\left[I_{R,k}\right] = \sum_i p_i \tag{5}$$

$$\begin{cases} E\left[O_{R,k+1} \middle| Q_{R,k} = 0\right] = 1 - \prod_i (1 - p_i) \\ E\left[O_{R,k+1} \middle| Q_{R,k} \geq 1\right] = 1 \end{cases} \tag{6}$$

Equ.6 figures out that the current time output of the router only depends on its queue length at the previous time step.

At the beginning, total packets input of the router is less than its maximum output, all the incoming packets will be forwarded real time. So, the system appears linearly and

$Q_R = 0$. However, with the packet sent probability increasing, input packets become greater than the maximum output capacity. Router shows nonlinear characteristics between its input and output, which means packets begin to queue in its buffer and Q_R increasing.

There can be a little interval Δt, during which all the source nodes have the same unchanged packet sent probability $p_{\Delta t}$. The total increase of Q_R is

$$Q_{R,\Delta t} = \max\left\{\sum\left(E\left[O_{R,\Delta t}\right] - E\left[I_{R,\Delta t}\right]\right), 0\right\} \tag{7}$$

When the packet queue length grows beyond the buffer capacity, the router will overflow, resulting in packets dropped. And the lifetimes of the rest packets will increase because of queuing in the buffer.

The nonlinear relation between input and output described above makes the router congested, which further affects packet forwarding of surrounding routers. In other words, the congestion will diffuse as packet transmission over the whole network. Finally, the entire network will be too congested to work efficiently.

In the actual World Wide Web (WWW) network, each webpage has different accessed frequency. The majority of user accesses gather in a few websites with high in-degree, which makes the network more vulnerable to the negative impact of nonlinear. Although it is very normal to improve network congestion by increasing trunk link bandwidth and enhancing router processing capacity in practice, the collapse of web servers still can not be avoid because of the nonlinear existing.

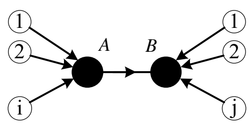

Fig. 3. Reducing Impact of nonlinear using distributed proxy

Fig.3 shows an extreme example without impact of nonlinear using distributed proxy. Nodes $1,\cdots,i,j$ send packets as described before. When A works as a router, $Q_A \neq 0$, B processes the incoming packet, $Q_B = 0$. If A is deployed to be proxy of B, it is obvious that $Q_A = 0$. So, we can see that the network can eliminate the negative impact of nonlinear partly by appropriate distributed proxy.

In the previous novel two-dimensional coupling model, the proxy approach can be explained in Fig.4. Node B in shadow grid 9 processes all the user access packets from the whole network, originally. Then, node A in shadow grid 5 is deployed as a distributed proxy of B. According to the routing strategy described in section 2, all user access packets from shadow grids 1, 2, 4 and part of packets from 3, 5, 7 will be redirected to the proxy node A. Packets from 6, 8, 9 will still be processed by node B. The

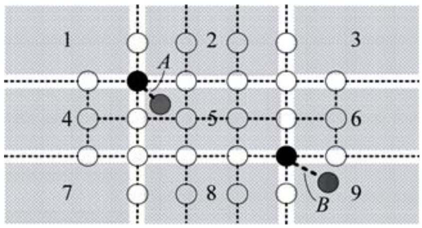

Fig. 4. Modified Two-dimensional coupling network model

deployed proxy node A, not only reduces the queue length of partial routers, but also makes all the user packets, which are original routed to B, spatially balanced in the entire network.

Given the cost constraints and other factors in actual network, it is unrealistic to eliminate the negative impact of nonlinear to the network performance completely. However, it is very feasible to reduce this kind of impact by deploying proxy node appropriately. Based on the novel two-dimensional coupling model, distributed proxy approach will be taken on the simulation verification in section 4.

4 Simulation Results and Analysis

The analysis of actual traffic in Tsinghua CERNET reveals the fact that 10% source nodes attracts 80% user access in WWW network, shown in fig.5.

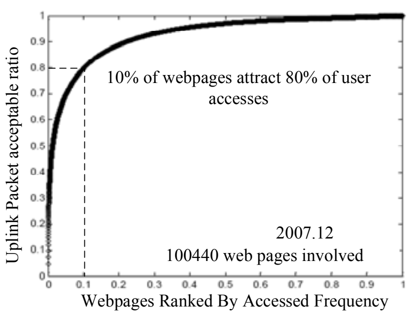

Fig. 5. Clustering Characteristic of User Access Behavior

That is the clustering characteristic of user access behaviors in actual network. The following simulations compare the performance changing of actual network, regular network and network with distributed proxy. In regular network, each node has the same accessed probability. The network model sizes are $L = 10, 20$. In each shadow grid there are 15 users or source providers. After 1000 time steps' evolution, the results are shown in figures 6, 7 and 8. Vertical solid lines indicate the phase transition critical, corresponding to the critical packets injected probability.

Fig. 6 and 7 are statistical results of queue length and packet lifetime, respectively. Clustering characteristic of user access behavior aggravates the negative affects of nonlinear to network, which arise to left-shift of phase transition critical, means the actual network more congested than the regular network.

The performance deterioration makes the network can not accept so many user accesses. Packets are dropped due to overflow of router buffer. In other words, the websites won't be capable of dealing with so many user access demands, and users can not receive requested information from the server. It is very difficult to guarantee the quality of services (QoS).

The power-law distribution of node degree in WWW network means there are a few nodes with high degree, called rich node here. According to the statistical ranked results shown in fig.5, the top 5% rich nodes are selected and deployed $c\lg k$ ($0 \le c \le 1$, k represents the node degree) distributed proxies at random positions in the model respectively. To some extent, these distributed proxy nodes balances the network traffic load, leading to the increasing of network critical injection rate, as shown in fig.6, 7 and 8. Right-shift of the phase transition critical illustrates an obvious performance improvement.

In this case, packets are redirected to the nearest proxy, with routing path shortened. Packets, originally gathered in one backbone trunk, are processed by several different routers connected to proxies. The whole network makes full use of buffer and process capacity of proxies, reducing buffer overflows and packet loss. The faster packets routed to the destination, the fewer life time they end with. Obviously, network performance has been greatly improved.

The routing distribution of data packets influences the balance of overall network traffic. With packet destination selected randomly, the network traffic appears balanced. Unfortunately, clustering characteristic of access behavior breaks up this balance in actual internet, leading to several backbone routers overloaded and inefficient packet transmission. Li et al. [12] studied the network security from the perspective of overall network characteristics, and figured it out when user accesses some important nodes not for useful information but attacking, network collapses rapidly. Distributed proxy approach can improve network robustness by reducing the risk of major nodes being attacked. Even this happens, the network services can be reconstructed and recovered rapidly.

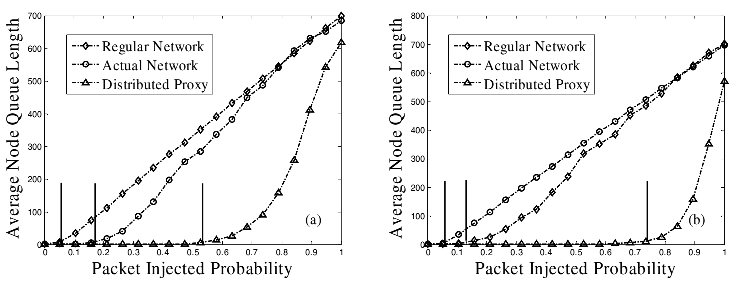

Fig. 6. Average node queue length statistical results in different network size. (a):L=10 (b):L=20.

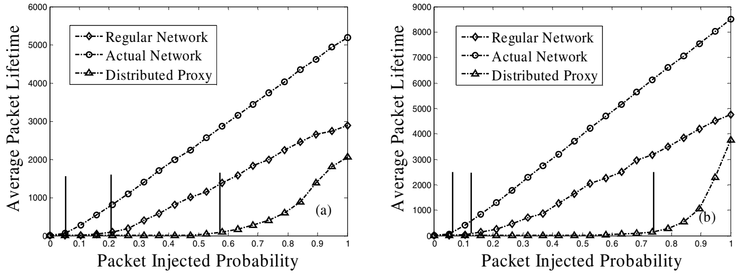

Fig. 7. Average packet lifetime statistical results in different network size. (a):L=10 (b):L=20.

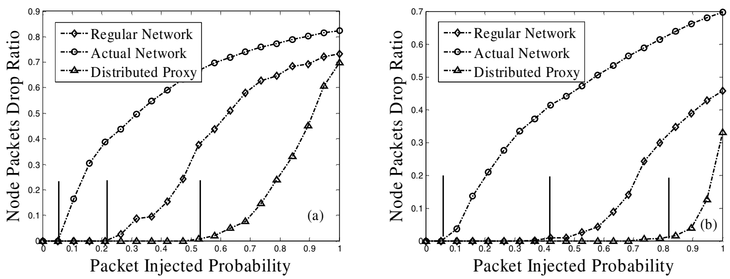

Fig. 8. Node packets dropped ratio statistical results in different network size. (a):L=10 (b):L=20.

5 Conclusions

In this paper, we analyze the nonlinear relationship between input and output of router node in detail, reveal the nonlinear mechanism of network performance phase transition and propose the distributed proxy approach to improve the network performance. Through theoretical analysis and simulation verification based on the novel two-dimensional coupling network model, we conclude as follows: 1) the queue length of router, packet life time and packet dropped ratio all arise phase transition with nonlinear existing; 2) compared to packet destination selected randomly in regular network, the extreme clustering characteristic of access behaviors in WWW networks gives arise to left-shift of phase transition critical, demonstrating deterioration of internet performance; 3) to solve this problem, distributed proxy approach is proposed to deploy several proxies, shortens the packet routing path. As a result, right-shift of the phase transition critical illustrates an improved performance of the network.

This paper is an important part of series studies. Research on various types of network phase transitions and the relationship to measure of proxies, the quantitative characterization of network nonlinear, optimization of network modeling and the bound of network performance, are all challenging research issues.

References

1. Willingers, W., Taqqus, M.S., Shennans, R., Wilson, D.V.: Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. IEEE/ACM Transactions on Networking 5(1), 71–86 (1997)
2. Csabai, I.: 1/f noise in computer network traffic. Physica A 27, 417–421 (1994)
3. Takayasu, M., Fukuda, K., Takayasu, H.: Application of statistical physics to the Internet traffics. Physica A 274, 140–148 (1999)
4. Hu, M.-B., Wang, W.-X., Jiang, R., Wu, Q.-S., Wu, Y.-H.: Phase transition and hysteresis in scale-free network traffic. Phys. Rev. E 75, 036102 (2007)
5. Tretyakov, A.Y., Takayasu, M., Takayasu, H.: Phase transition in a computer network model. Physica A 253, 315–322 (1998)
6. Ohira, T., Sawatari, R.: Phase transition in a computer network traffic model. Phys. Rev. E 58, 193–195 (1998)
7. Fuks, H., Lawniczak, A.T.: Performance of data networks with random links. Mathematics and Computers in Simulation 51, 101–117 (1999)
8. Jian, Y., Yong, R., Feng, L., Xiu-Ming, S.: Phase transition and collective correlation behavior in the complex computer network. Acta Phys. Sin. 50, 1221–1225 (2001)
9. Feng, L., Xiu-Ming, S., Yong, R.: Long-range correlation in computer network. Acta Phys. Sin. 53, 373–378 (2004)
10. Feng, L., Yong, R., Xiu-Ming, S.: A simple cellular automata model for packet transport in the Internet. Acta Phys. Sin. 51, 1175–1180 (2002)
11. Lei, W., Shu-Hua, Z., Jian, Y., Yong, R., Xiu-Ming, S.: Influence of virtual networks to internet collective behavior. Acta Phys. Sin. 56, 36–42 (2007)
12. Ying, L., Xiu-Ming, S., Yong, R.: Dynamic properties of epidemic spreading on finite size complex networks. Chinese Physics 14, 2153–2157 (2005)