2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Telephone Traffic Analysis Based on a Scale-Free User Network

Yongxiang Xia, Chi K. Tse, Wai Man Tam, Francis C. M. Lau, and Michael Small

Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong, China
http://chaos.eie.polyu.edu.hk

Abstract—In this paper we analyze the traffic of telephone systems. Unlike classical traffic analysis where users are assumed to be connected uniformly, our proposed method employs a small-world scale-free network to model the behavior of telephone users. Each user has a fixed set of acquaintances with whom the user may communicate, and the number of acquaintances follows a power-law distribution. We show that network traffic is greatly influenced by the user network behavior, and that call blocking probability is generally higher in the case of a scale-free user network. It is also shown that the carried traffic intensity is practically limited not only by the network capacity but also by the scale-free property of the user network.

1. Introduction

Models for traffic analysis have been derived by fitting the existing traffic data under particular sets of conditions [1]. Since the underlying mechanisms of the complex network behavior are unknown or simply not taken into account in the modeling process, such models fall short of a clear connection with the actual physical processes that are responsible for the behavior observed in the traffic data.

Recent study of small-world and scale-free properties of so-called *complex networks* has motivated research in the modeling of practical networks based upon certain specific network topologies that possess properties closely resembling those of realistic physical networks [2, 3]. In general terms, a *complex network* may be characterized by a large number of nodes and a set of complex relationships between them [4, 5]. From the viewpoint of complex networks, the user network underlying any communication network exhibits small-world and scale-free properties [3]. Up to now, complex network behavior in telephone systems has been rarely considered. Aiello *et al.* [6] studied the scale-free property in the daily traffic of a long-distance telephone call graph. However, to date, practical traffic analysis based upon a small-world and scale-free user network is still unavailable. A relevant preliminary work can be found in Xia *et al.* [7].

In this paper we attempt to incorporate a *user network model* in the analysis of telephone system traffic. We aim to provide a clear connection between the user network behavior and the network traffic, and illustrate how network traffic data can be more realistically simulated with the inclusion of a proper user network behaviorial model. This study clears up several misconceptions. Telephone traffic (including mobile network traffic) cannot be considered without taking into account the way in which human users are connected in the real world. The fact that human networks possess small-world and scale-free properties can change the way network resources have to be planned. For instance, we will show that limited network capacity is not always the cause of call blockings while the scale-free property of the user network may be the root of the problem.

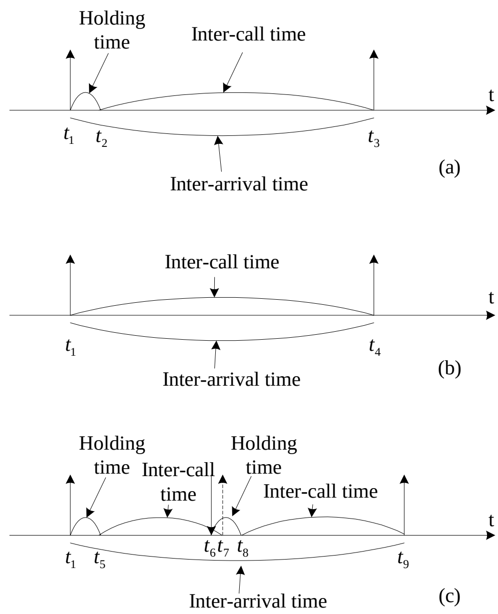

Figure 1: Three typical calling processes. (a) Call established; (b) call blocked; (c) call cancelled.

2. Traffic Analysis in Telephone Systems

In a telephone system, "traffic" refers to the accumulated number of communication channels occupied by all users. Different from the user network, the telephone system is a directed complex network, in which each edge has a direction from the caller to the receiver. For each user, the call arrivals can be divided into two categories: *incoming calls* and *outgoing calls*. Here, incoming calls refer to those being received by a user, and outgoing calls refer to those being initiated by that user. Since every incoming call for one user must be originated from an outgoing call of another user, we only need to consider outgoing calls from each user when we analyze the network traffic. If not specified, the term *call arrival* means *outgoing call arrival* in the rest of this paper.

Outgoing calls are initiated randomly. If a call arrives and the conversation is successfully established, both the caller and the receiver will be engaged for a certain duration commonly known as *holding time*. The length of the holding time is also a random variable. Thus, the traffic load depends on the rate of call arrivals and the holding time for each call [8, 9]. Figure 1 shows three cases of the calling process.

Case I: Call established. When an outgoing call arrives at time t_1, a receiver is randomly selected. If this receiver is idle at that time and the network has an idling channel, a call is successfully established. The caller and receiver will be engaged for a duration of holding time ($t_2 - t_1$). The call ends at time t_2. The *inter-call time* ($t_3 - t_2$) is the duration between the end of this call and the beginning of the next outgoing call arrival. Also, the inter-arrival time is equal to the sum of the holding time and the inter-call time, as depicted in Fig. 1 (a).

Case II: Call blocked. Suppose two outgoing calls are made at t_1 and t_4. A call blocking may occur due to two reasons. First, if the receiver is engaged with another call at time t_1, then any new call attempt will fail and a call blocking is said to occur. Another reason for call blockings is the limited network capacity. If all channels are occupied at time t_1, the new call attempt is blocked. The telephone system is usually considered as a "lossy" system, in which the blocked call simply "disappears" from the network. In this case the inter-arrival time is equal to the inter-call time (i.e., $t_4 - t_1$), as shown in Fig. 1 (b).

Case III: Call cancelled. In this case, an outgoing call is supposed to take place at time t_7. However, if an incoming call has arrived ahead of it and the conversation is still going on at time t_7, the outgoing call attempt will be cancelled. Since this call attempt has not been initiated, it is counted as neither a call arrival nor a call blocking. When the conversation ends at time t_8, another inter-call time is assumed before the next outgoing call arrives at time t_9. In this case, the inter-arrival time is ($t_9 - t_1$), as illustrated in Fig. 1 (c). Of course, at time t_9, it is also possible that the user is engaged with another call. Then, the call arrival at time t_9 will again be cancelled, and the inter-arrival time will be extended accordingly.

Clearly, the shortest inter-arrival time is equal to the inter-call time, which happens only in Case II. In our subsequent analysis, the above three cases of the call arrival process will be considered.

The holding time and the inter-call time are usually modelled by some random variables with exponential distribution. The probability density function (PDF) of the holding time is given by

$$f_1(t) = \frac{1}{t_m} e^{-t/t_m} \tag{1}$$

where t_m is the average holding time. The PDF of the inter-call time is given by

$$f_2(t) = \mu_i e^{-\mu_i t} \tag{2}$$

where $1/\mu_i$ is the average inter-call time. The holding times of all users have the same distribution, but the mean values of the inter-call times for different users may be different.

As shown in Fig. 1, the inter-arrival times for the three cases are different. However, if we examine the traffic over a sufficiently long period of time (e.g., 60 min), we can obtain the average call arrival rate λ_i, which is the average number of call arrivals per unit time, for user i. Thus, the average arrival rate for the whole network is

$$\lambda = \sum_{i=1}^{N} \lambda_i \tag{3}$$

where N is the total number of users in the network.

A commonly used measure of traffic is the *traffic intensity* [1], which is defined by

$$A = 2\lambda t_m, \tag{4}$$

and represents the average traffic offered over a period of time. It is dimensionless, but is customarily expressed in units of erlang. Notice that there is a factor of 2 in (4), which arises from the fact that both the caller and the receiver stay in the same telephone system. Thus, two channels are used for each call conversation.

In a telephone system, the *offered traffic* is the total traffic that is being requested by users, and the carried traffic is the actual traffic being carried by the network, which can be found as the sum of the holding times of all call conversations. In practice, due to limited network capacity and some user behavior, a certain percentage of the offered traffic experiences *call blocking.* Hence, the carried traffic is

$$A_{\text{carried}} = A_{\text{offered}} \times (1 - p_{\text{blocking}}) = 2\lambda t_m (1 - p_{\text{blocking}}) \tag{5}$$

where A_{carried} and A_{offered} denote the carried traffic and the offered traffic, respectively, and p_{blocking} represents the blocking probability of a call.

The telephone system is typically measured in terms of the average activity during the busiest hour of a day. During the busiest hour, the average contribution of one user to the traffic load is typically between 0.05 and 0.1 erlang. The average holding time is 3 to 4 min [1].

3. Effect of User Network

Formally, we may describe a user network of a telephone system in terms of nodes and connections. A node is a user, and a connection between two nodes indicates a possibility that these two users may call each other. In reality, people usually only call their own acquaintances, such as family members, colleagues and friends. Although a person may call other people which are not his acquaintance in practice, the probability is so lower that we can safely ignore it. Thus, in a user network, a connection connects a pair of acquaintances. In this paper we consider two kinds of user network, i.e., uniform user network and scale-free user network.

In a uniform user network, each user has a fixed number of acquaintances, $n_i = \overline{n}$ for all i, and $\overline{n} \ll N$. In such a user network, the effect of each user is identical.

In a small-world scale-free user network, the number of acquaintances for different users may be different. If a user has more acquaintances, the probability of him making/receiving a call at any time is higher. For user i, the number of acquaintances n_i is a random number. It has been found that many human networks are small-world scale-free networks, with n_i typically following a power-law distribution [3]:

$$p(n_i) \sim n_i^{-\gamma} \tag{6}$$

where $p(n_i)$ is the probability that user i has n_i acquaintances and $\gamma > 0$ is the characteristic exponent. Figure 2 shows a power-law distribution of n_i according to Eq. (6). We clearly see that a relatively small number of users have a large number of acquaintances.

Since the probability that a user with more acquaintances makes/receives calls is higher, the mean value of this user's inter-call time is smaller. In order to show this inequality, we assume

$$\mu_i = p_0 n_i, \tag{7}$$

where p_0 is a constant of proportionality. This relation is valid for the uniform user network and the scale-free user network. The only difference is that $n_i = \overline{n}$ is a fixed number in the uniform network, whereas n_i is a random variable in the scale-free network. For a fair comparison, we set the expectation of n_i in the scale-free user network to be equal to $\overline{n}$ in the following simulation. In this way, $E\{\mu_i\}$ for both user network configurations are identical.

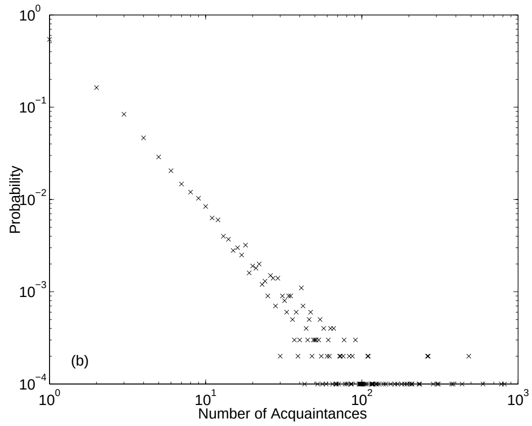

Figure 2: Scale-free user network. Probability of user i having n_i acquaintances versus n_i, showing power-law distribution. Mean n_i is 5.

4. Simulation Results

We consider a telephone system of N users. Users are located in M subsystems, each supporting N/M users. The use of subsystems is to reflect the actual case in modern telephone systems. In a fixed telephone system, the subsystems are the central offices; in a cellular mobile system, the subsystems are referred to as cells. Here, for simplicity, we assume that users remain in their subsystems for the entire period of simulation. Two user network configurations, namely, the uniform network and scale-free network, are considered.

In the user network, each user has his own acquaintance list. The following two-step method is used to construct the scale-free user network. First, the number of acquaintances n_i for user i is determined. In the uniform user network, n_i is fixed to $\overline{n}$. On the other hand, in the scale-free user network, n_i is a power-law-distributed random variable with the expectation $\overline{n}$. Next, the acquaintance lists are filled by randomly selecting acquaintances. The relationship of acquaintance is bi-directional. If user i is selected as an acquaintance of user j, then user j is automatically added into user i's acquaintance list. When a user is going to make a call, he randomly chooses a receiver from his acquaintance list.

The simulation parameters are set as follows:

$N = 10000$, $M = 4$, $\overline{n}$ = average $n_i = 5$,
$p_0 = 1/500$ call/min/acquaintance,
$t_m = 4$ min.

From the above setting, we can calculate the average inter-call time $1/\overline{\mu} = 100$ min for both user networks.

Figure 3(a) shows the blocking probability versus the channel capacity (i.e., the number of channels provided in each subsystem). When the channel capacity is very limited, almost all call arrivals are blocked. As the channel capacity increases, some of the arrived calls are successfully set up. The call blocking probability drops. However, as the capacity reaches a certain threshold, the blocking probability settles to a constant value. This clearly shows that when the channel capacity is beyond the threshold, channel capacity is no longer a factor for call blockings and user engagement becomes the only limiting factor. Further, the channel capacity threshold is related to the user network configuration. Our simulation (for the chosen set of parameters) shows that the capacity threshold for the uniform user network is about 210 channels per subsystem, and is only about 100 channels per subsystem for the scale-free user network. Moreover, the blocking probability for the scale-free user network settles to about 44%, and is much higher than that for the uniform user network, which is about 7%. The generally higher blocking probability for the scale-free user network is caused by call concentration on a small number of users who have a relatively large number of acquaintances.

Figure 3(b) shows the actual call arrival rate versus the channel capacity. From this figure, we can make two main observations. First, the threshold effect exists in both user network configurations. Before the capacity reaches the threshold, the call arrival rate decreases as channel capacity increases. When the channel capacity reaches the threshold, the call arrival rate is almost fixed. The small fluctuations in the resulting curves are due to the randomness of call processes in our simulation. Second, noticeable differences between the simulation results of the two user networks are found. The call arrival rate for the scale-free user network declines more rapidly than that for the uniform user network. Furthermore, the thresholds for the two user networks are different, and the call arrival rates beyond the corresponding thresholds are also different.

The decrease of the call arrival rate as channel capacity increases is due to the complex calling processes. The average inter-arrival times in the three cases are different. The shortest inter-arrival time happens in Case II. The actual calling process is the combination of the three typical calling processes. When the channel capacity is low, the channels are more likely to be fully occupied and Case II (i.e., call blocking) is more likely to occur. The average inter-arrival time is thus shorter, and the average arrival rate is higher. As the channel capacity increases, the blocking probability drops. Thus, the average inter-arrival time becomes longer, making the average arrival rate lower. When the channel capacity reaches the threshold, the blocking probability becomes steady, and the average call arrival rate remains almost unchanged.

The resulting carried traffic intensities are shown in Fig. 3(c). The carried traffic intensity is a function of the call arrival rate and blocking probability, as in (5). Hence, when a drop in call arrival rate is "over-compensated" by a reduction in blocking probability, the net effect is an increase in carried traffic intensity. This phenomenon occurs when the channel capacity is increased initially. As the channel capacity is increased beyond the threshold, both the call arrival rate and the blocking probability arrive at constant values, and the corresponding carried traffic intensity also becomes steady.

The simulated results may seem to deviate from our usual expectation. The normal way to avoid call blockings is to increase the system capacity. But our simulation results show that in addition to inadequate channel capacity, the user network configuration has a profound influence on call blockings. Increasing the system capacity may not solve the problem. The user network configuration must be considered when making telephone system planning. Our simulation also shows that the traffic for the scale-free user network differs significantly from that for the uniform user network, which is usually assumed in classical traffic analyses. Because of the scale-free nature of human networks, analyses based on a scale-free user network should reflect more realistic traffic scenarios.

In the scale-free network, γ is a very important parameter. As shown in (6), a smaller γ corresponds to a gentler slope of the

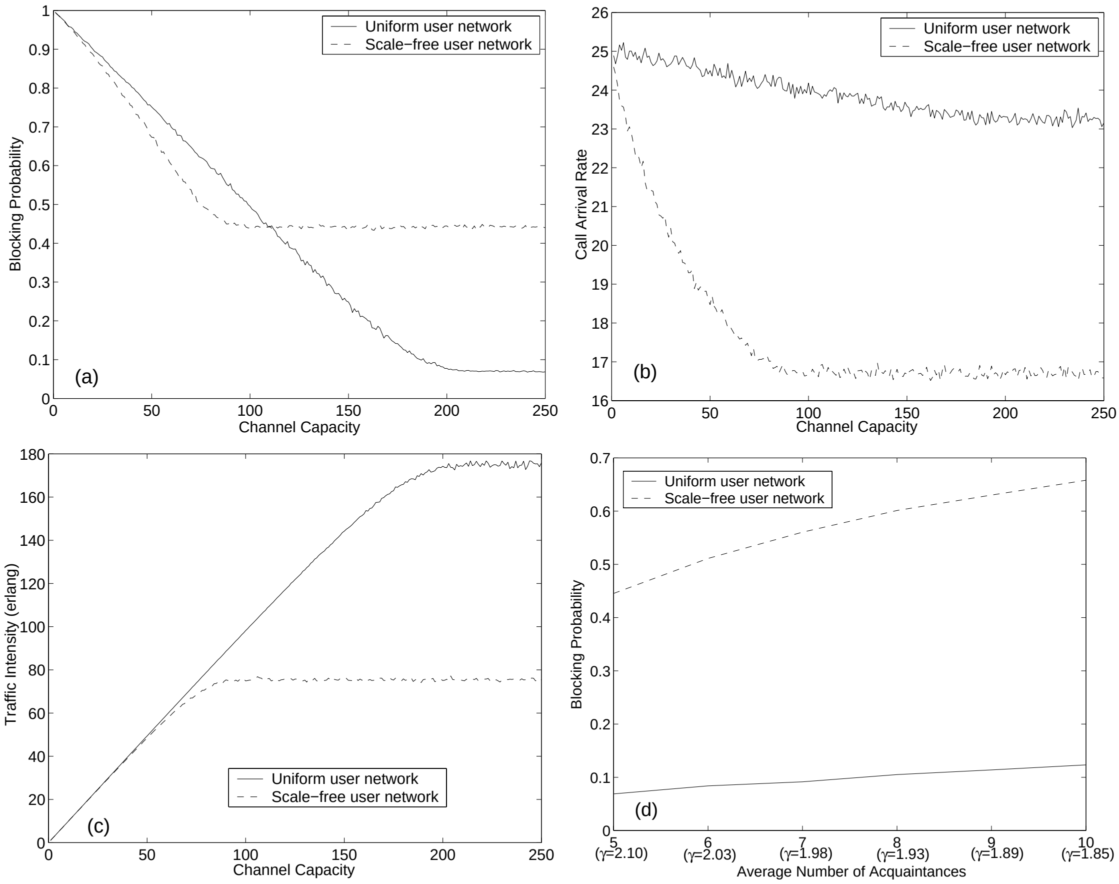

Figure 3: Simulation results. (a) Blocking probability versus channel capacity; (b) average call arrival rate versus channel capacity; (c) carried traffic intensity versus channel capacity; (d) blocking probability versus average number of acquaintances.

power-law distribution, which means that more users have a large number of acquaintances. Hence, the average number of acquaintances $\overline{n}$ increases as γ decreases. Figure 3(d) shows the blocking probability with $\overline{n}$ changes. Corresponding values of γ are also marked in the figure. In order to focus on the effect of $\overline{n}$, we eliminate the effects of the channel capacity by setting it to infinity in the simulation. The figure clearly indicates that comparing the telephone system with a uniform user network, the system with a scale-free user network has a much higher blocking probability.

5. Conclusions

This paper studies the telephone system traffic from a scale-free user network perspective. Two major factors, namely, the channel capacity and user network configuration, are identified as being pivotal to call blockings. Simulation results show that the traffic assuming a scale-free user network differs substantially from the traffic assuming a uniform user network. Our final conclusions are that telephone system traffic is greatly influenced by user behavior, and that beyond a certain capacity threshold call blockings are not likely to be reduced by increasing channel capacity (adding extra resources or intensifying investments) which would have been the usual expectation. Thus, a clear, though obvious, lesson to be learnt from this traffic analysis is that any strategy for altering the traffic in any manner must take into account the scale-free property of user networks.

References

[1] J. C. Bellamy, *Digital Telephony,* New York: Wiley, 2000.

[2] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature,* vol. 393, pp. 440–442, June 1998.

[3] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics,* vol. 74, pp. 47–97, 2002.

[4] X. F. Wang and G. Chen, "Complex networks: small-world, scale-free and beyond," *IEEE Circuits & Systems Magazine,* vol. 3, no.1, pp. 6–20, 2003.

[5] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review,* vol. 45, no. 2, pp. 167–256, 2003.

[6] W. Aiello, F. Chung, and L. Lu, "A random graph model for massive graphs," *Proc. 32nd Annual ACM Symposium on Theory of Computing,* pp. 171–180, 2000.

[7] Y. Xia, C. K. Tse, W. M. Tam, F. C. M. Lau and M. Small, "Scale-free user network approach to telephone network traffic analysis," *Physical Review E,* to appear.

[8] Siemens Aktiengesellschaft, *Telephone Traffic Theory: Tables and Charts,* Berlin: Siemens, 1970.

[9] R. A. Thompson, *Telephone Switching Systems,* Boston: Artech House, 2000.