Modeling the Telephone Call Network

Wai M. Tam, Francis C. M. Lau and Chi K. Tse
Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong
http://chaos.eie.polyu.edu.hk

***Abstract*— It is natural to envisage that in a telephone network, some telephone numbers originate or receive more calls than others. Indeed, real-life data have verified the conjecture that the number of calls originated from or received by a telephone number in a network has a power-law property. Further, the number of calls made or received by the same telephone number can be very different. In this paper, we construct a self-growing complex network for modeling the aforementioned telephone call network. The complex network obtained is a directed and weighted network. Moreover, the nodes (telephone numbers) of the network exhibit power-law properties in the following aspects: total-degree, out-degree, in-degree, total-strength, out-strength and in-strength.**

I. INTRODUCTION

In the past decade, a large number of networks, such as the Internet [1], [2], the social networks [3], [4] and the transportation networks [5], [6], were found to exhibit scale-free property. Early studies of such networks focused on the un-weighted scenario, in which all edges had a weight of 0 or 1, regardless of the interaction frequency between two connected nodes. In many occasions, the un-weighted network was not sufficient in characterizing the actual network.

Later, researchers shifted their attention to a more realistic model—the weighted network. Two typical weighted networks, namely the world-wide airport network (WAN) and the scientist collaboration network (SCN), had been widely studied [7], [8], [9], [10]. Based on the statistical data of 128 airports in China, it was found that over a week, the number of flight seats between two airports were almost the same in both directions [5]. Thus, the WAN could be modeled as a complex network in which a node denoted an airport and the weight of an edge represented the number of seats provided by the flights between the two airports. As for the SCN model, each node stood for a scientist and the weight of an edge represented the number of coauthored papers between two scientists. Note that in the modeling of both WAN and SCN, it is sufficient to connect the nodes with edges that have weights but no direction, i.e., undirected.

In the case of a telephone call network (TCN) model, each edge should be characterized by two parameters, namely weight and direction, which specify the frequency of calls and the telephone number initiating such calls, respectively. Real-life data have also shown that the number of calls originated from or received by a telephone number in a network has a power-law property [11]. In other words, a few telephone numbers make or receive lots of calls while a vast majority of telephone numbers make or receive very few calls. Further, the number of calls made or received by the same telephone number can be very different. In order to perform an accurate investigation into the call dynamics of the telephone network, it is necessary to produce a more precise model for the call behavior of the telephone numbers. This is particularly useful in predicting the dynamics of the telephone calls during incidents like earthquakes or typhoons, when there is a sudden surge in the number of telephone calls. In this paper, we construct a self-growing complex network for modeling the TCN reported in [11]. The complex network obtained is a directed and weighted network. Moreover, the nodes (telephone numbers) of the network have power-law properties in the following aspects: total-degree, out-degree, in-degree, total-strength, out-strength and in-strength.

Before we proceed, we define the variables used in this paper. We denote the weight of an edge (the number of calls per unit time) directed from node i to node j by w_{ij}. (Note that w_{ij} and w_{ji} are not equal in general.) Also, the out-strength, in-strength and total-strength of node i are defined as $\hat{s}_i = \sum_j w_{ij}$, $\check{s}_i = \sum_j w_{ji}$ and $s_i = \hat{s}_i + \check{s}_i$, respectively. Further, the out-degree $\hat{k}_i$ and in-degree $\check{k}_i$ of node i are defined as the number of outgoing edges and incoming edges of node i. Finally, the total-degree k of node i is simply the total number of neighboring nodes connected to node i. But note that it does not equal to the sum of out-degree and in-degree in general.

The project is supported by Hong Kong Polytechnic University under project 1-BBZA.

II. MODEL CONSTRUCTION

In this section, we propose a network construction model to generate a directed, weighted complex network with similar characteristics as a telephone call network. The nodes of the network represent the telephone numbers while the weights of each edge correspond to the frequencies that calls are made from one telephone number to another.

A. Initialization

We start with a directed network with N_0 fully-connected nodes (telephone numbers), i.e., all N_0 nodes are interconnected by a pair of inverse edges (directed edges of opposite directions). Moreover, the weight of each directed edge is w_0 (number of calls per unit time), i.e., $w_{ij} = w_{ji} = w_0$ for all i and j. One example with $N_0 = 3$ is shown in Fig. 1.

B. Creation of New Nodes

At each time step, a new node, with node number i, is created. It is then connected to m existing nodes by purely outgoing, purely incoming or both outgoing and incoming

1-4244-0921-7/07 $25.00 © 2007 IEEE.

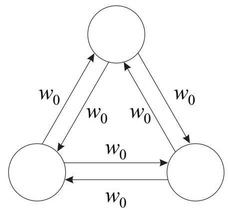

Fig. 1. A directed and fully-connected network with 3 nodes.

edges. We apply a preferential treatment to attach the new node to existing nodes, i.e., the probability that the new node i is attached to an existing node j is given by

$$\Pi_j = \frac{s_j}{\sum_u s_u}. \quad (1)$$

This rule describes the scenario that an existing node (telephone number) which has a larger strength (total number of calls) is more likely to be attached to (make contact with) the new node .

Suppose an existing node j is selected. If node j has only outgoing edges (incoming edges), then a directed edge is added from node j to node i (node i to node j), and the weight of the new edge is assigned with unity, i.e., $w_{ji} = 1$ ($w_{ij} = 1$). These cases may correspond to telephone booths and emergency telephone numbers for which only outgoing calls and incoming calls, respectively, are allowed.

For the case when node j contains both outgoing and incoming edges, we have two different scenarios.

- With a probability of p, a pair of inverse edges are added between nodes i and j. The weights of the edges are assigned with unity, i.e., $w_{ij} = w_{ji} = 1$. This rule is established to describe the situation when the two nodes i and j are acquaintances (relatives, friends or colleagues) of each other. Also, they have equal probabilities of calling each other.
- With a probability of $1 - p$, one directed edge is added between nodes i and j. Here, we make a general assumption that the probability of the edge in a particular direction (incoming or outgoing) is proportional to the strength (in-strength or out-strength) of the existing node along the same direction. That is to say, the probability that the edge goes from i to j is given by $\check{s}_j/s_j$, and the probability that the edge directs from j to i equals $\hat{s}_j/s_j$. In both cases, the weight of the new edge is equal to one. Also, under such an arrangement, a new node (telephone number) may evolve into having only outgoing or incoming edges, corresponding to a new telephone booth or emergency number.

C. Update of Edge Weights and Creation of New Edges Among Existing Nodes

During each time step, we also update the weights of the edges and create new edges among the existing nodes.

1) We assume that if a telephone number makes/receives calls frequently, the probability that this telephone number will make/receive more calls is high. That is to say, a busy telephone number will become busier as the network evolves. Therefore at each time step, we select n pairs of nodes based on their individual in-strength and out-strength.
For each of the n cases, according to the out-strengths of the existing nodes, a node v (caller) is first chosen with a probability of

$$\Pi_v = \frac{\hat{s}_v}{\sum_u \hat{s}_u}. \quad (2)$$

Then, based on the in-strengths of the remaining nodes, another node j (callee) is selected with a probability of

$$\Pi_j = \frac{\check{s}_j}{\sum_{u,u\neq v} \check{s}_u}. \quad (3)$$

If there exists a directed edge from node v to node j, the weight of the edge will be incremented by one, i.e., $w_{vj} = w_{vj} + 1$. Otherwise, a new edge with weight 1 will be added from node v to node j, i.e., $w_{vj} = 1$.
2) We also assume that all existing communication edges will be enhanced with time. Therefore, if there is a directed edge between two nodes, the weight of this edge should increase with time. Also, a edge with a larger weight will have a higher chance of being strengthened. At each time step, we pick l directed edges and increase their weights by one. The probability of the directed edge from node v to node j being selected is given by

$$\Pi_{v\to j} = \frac{w_{vj}}{\sum_z \sum_{u,u\neq z} w_{zu}}. \quad (4)$$

Note that when $p = 1$ and $n, l = 0$, our model degenerates to the BA model [13] and an un-directed, un-weighted complex-network model will be created, i.e., $w_{ij} = w_{ji} = 1$ and $\hat{k}_i = \check{k}_i$.

III. Results and Discussions

In this section, we present the characteristics of the complex networks created by our proposed method. Unless otherwise stated, the results are averaged over 10 independently generated networks. The parameters used to generate the networks are given as follows: $m = 3$, $w_0 = 1$, $N = 5000$ (total number of nodes in the final network) and $N_0 = \max(m,\ \lceil 1/2 + 1/2\sqrt{1+4n}\rceil),\ \lceil 1/2 + 1/2\sqrt{1+4l}\rceil)$. Note that N_0 has been selected in such a way that there is a sufficient number of distinct nodes and edges to choose from during the first node-attachment, weight-update and edge-creation process.

First, we present the results for $p = 0$. For this particular case, we have also derived analytically the probability distribution of the strengths of the nodes. Due to the space limitation, the details are not shown here. But the analysis has indicated that as the number of time steps increases, the probability distributions of the in-strength, out-strength and total-strength of the nodes all approach power-law with the same exponent. In other words, denoting the exponent by α, we have $P(\check{s}) \propto \check{s}^\alpha, P(\hat{s}) \propto \hat{s}^\alpha$ and $P(s) \propto s^\alpha$ as time goes to infinity. Moreover, our analysis produces an expression for α, which is

$$\alpha = 2 + \frac{m}{m + 2n + 2l}. \quad (5)$$

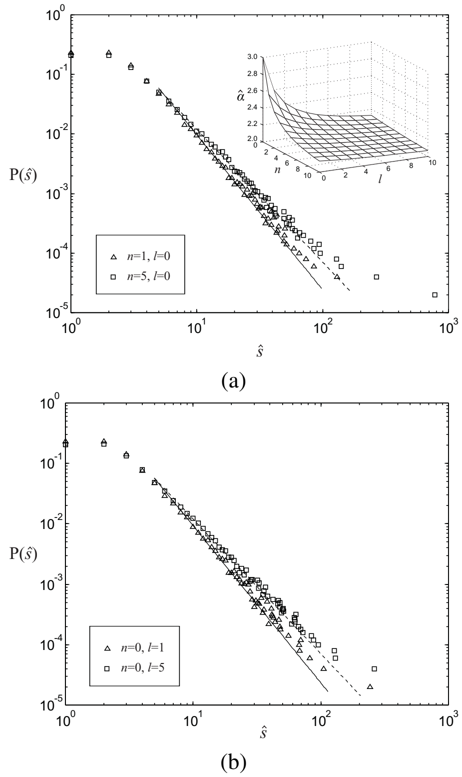

Fig. 2. Probability distribution of the out-strength $\hat{s}$ of the nodes when $p = 0$. For each set of parameters, the results are averaged over 10 independently generated networks. The inset shows the analytical values of α for $0 \leq n, l \leq 10$. (a) $l = 0$. The solid ($n = 1$) and the dashed ($n = 5$) lines are plotted using (5); (b) $n = 0$. The analytical results are given by the solid ($l = 1$) and the dashed ($l = 5$) lines.

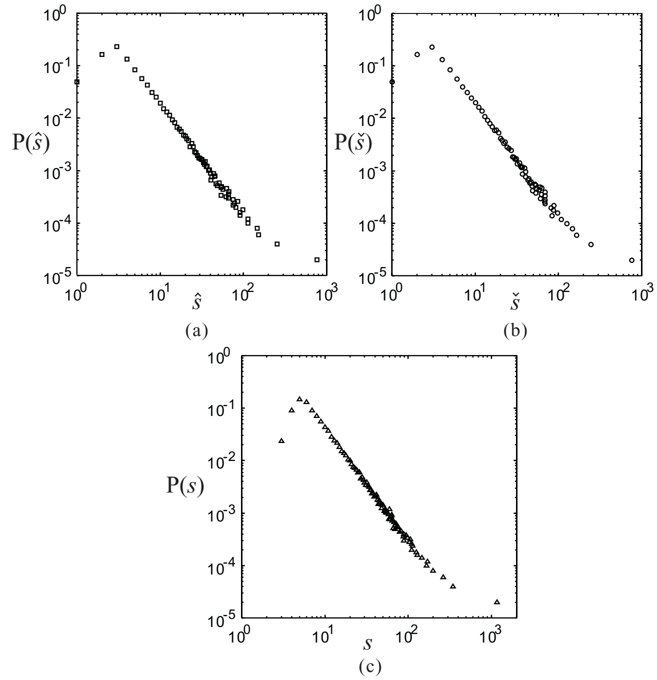

Fig. 3. Probability distribution of the (a) out-strength $\hat{s}$; (b) in-strength $\check{s}$; (c) total-strength s of the nodes. $p = 0.6$, $n = 3$ and $l = 3$. The results are averaged over 10 independently generated networks.

From the above equation, we observe that when n and l increase, the exponent decreases. Also, both n and l affect the exponent equally. Figure 2(a) and (b) plots the probability distribution of the out-strength of the nodes for $l = 0$ and $n = 0$. It can be observed that the distributions obey power-law and are consistent with the results derived analytically. As n or l increases, α reduces and tends to 2.0.

Next we investigate the case when $p = 0.6$, $n = 3$ and $l = 3$.

1) Figure 3 plots the probability distributions of the out-strength, in-strength and total-strength of the nodes. It is easily seen that the three distributions obey power-law with similar exponents.
2) In Fig. 4, we further plot the out-strength, in-strength and total-strength of the nodes versus out-degree, in-degree and total-degree, respectively. It can be clearly seen that the log-log graphs produce straight lines of similar slopes. We can therefore conclude that $s \propto k^{\beta}$ [12], $\hat{s} \propto \hat{k}^{\hat{\beta}}$ and $\check{s} \propto \check{k}^{\check{\beta}}$, where the three scaling exponents, denoted by $\hat{\beta}, \check{\beta}$ and β, are of similar values.
3) Because of the aforementioned results, it is clear that the probability distributions of the node degrees, as shown in Fig. 5, are also power-law distributed. It manifests that a large number of telephone numbers keep in touch with a few telephone numbers, while only a small number of telephone numbers have contact with lots of telephone numbers.

Finally, we investigate the weights of the edges of opposite directions, i.e., w_{ij} and w_{ji}. Figure 6 plots w_{ij} versus w_{ji} for two particular networks. It can be observed that the values of w_{ij} and w_{ji} are different in most cases. The results indicate that for any pair of telephone numbers, the number of calls may not be symmetrical in the two directions , i.e., one telephone number may make more calls to than receive calls from the other telephone number. We also observe that when p is increased from 0 to 0.6, the values of w_{ij} and w_{ji} are closer.

IV. Conclusions

In this paper, we have proposed a method to construct a self-growing complex network for modeling the telephone call network. The complex network obtained is a directed and weighted network. Further, the nodes (telephone numbers) of the network have scale-free properties in the following aspects: total-degree, out-degree, in-degree, total-strength, out-strength and in-strength. Moreover, the weights of the edges between two nodes are usually not identical. Such characteristics are consistent with the observations made on a real telephone call network (TCN) [11]. Having constructed a model for the TCN, researchers will be able to predict the call dynamics of the network more precisely when some un-expected events (such as earthquake or typhoon) occur, causing a sudden surge in the number of telephone calls. Measures can then be proposed to reduce the impact on the telephone network when such events do happen. For example, call attempts to telephone numbers with larger total-strengths will be blocked with a

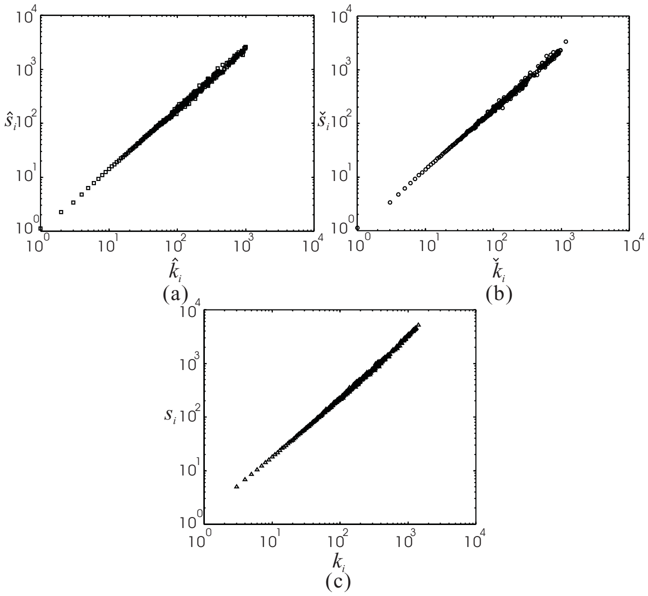

Fig. 4. Plot of strength versus degree. $p = 0.6$, $n = 3$ and $l = 3$. The results are averaged over 10 independent networks for each case. (a) Out-strength $\hat{s}$ versus out-degree $\hat{k}$; (b) in-strength $\check{s}$ versus in-degree $\check{k}$; (c) total-strength s versus total-degree k.

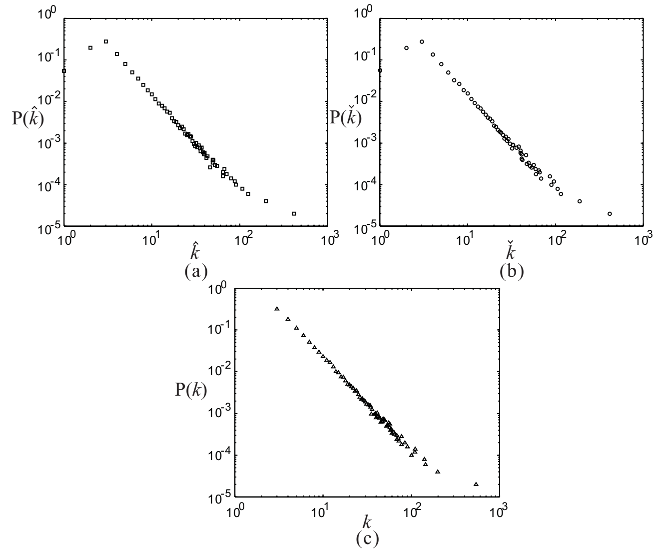

Fig. 5. Probability distribution of the (a) out-degree $\hat{k}$; (b) in-degree $\check{k}$; and (c) total-degree k of the nodes. $p = 0.6$, $n = 3$ and $l = 3$. The results are averaged over 10 independent networks.

higher probability because it is very likely that those telephone numbers are already engaged. Such earlier call blocks will alleviate the loading of other components of the telephone network, enhancing the chance of other telephone calls to get through. The effect of such call blocking mechanisms will be further investigated in future.

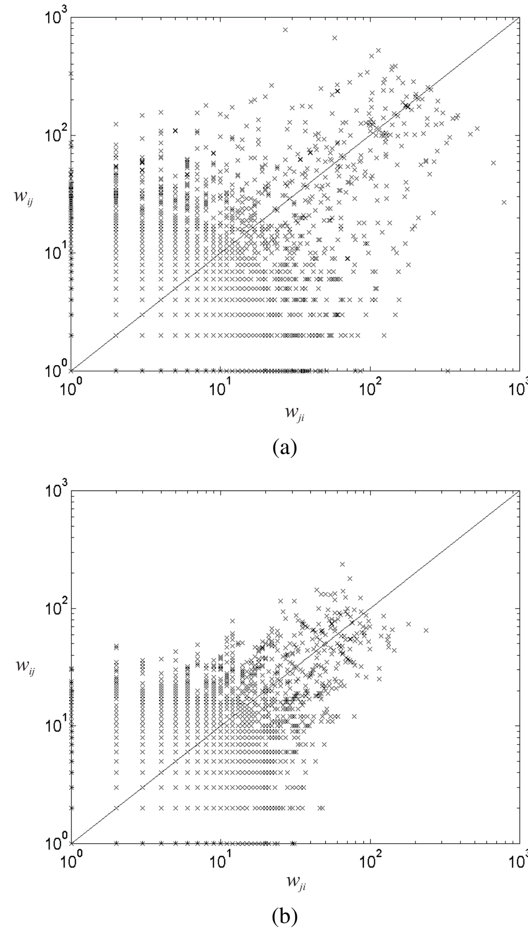

Fig. 6. Plot of w_{ij} versus w_{ji} for two particular networks. $n = 3$ and $l = 3$. (a) $p = 0$; (b) $p = 0.6$.

REFERENCES

[1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," *ACM SIGCOMM Computer Communication Review*, vol. 29, issue 4, pp. 251–262, October 1999.
[2] R. Govindan and H. Tangmunarunkit, "Heuristics for Internet map discovery," *Proc. IEEE INFOCOM 2000*, Tel Aviv, Israel, vol. 3, 2000, pp. 1371–1380.
[3] S. Redner, "How popular is your paper? An empirical study of the citation distribution," *The European Physical Journal B*, vol. 4, pp. 131–134, 1998.
[4] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.
[5] W. Li and X. Cai, "Statistical analysis of airport network of China," *Phys. Rev. E*, vol. 69, 046106-1–6, 2004.
[6] G. Bagler, "Analysis of the airport network of India as a complex weighted network," *eprint cond-matter*, paper number 0409773, 2004.
[7] M. E. J. Newman, "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality," *Phys. Rev. E*, vol. 64, 016132-1–7, 2001.
[8] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks," *Proc. Natl. Acad. Sci.*, vol. 101, pp. 3747–3752, March 2004.
[9] W. Wang, B. Hu, T. Zhou, B. Wang, and Y. Xie, "Mutual selection model for weighted networks," *Phys. Rev. E*, vol. 72, 046140-1–8, 2005.
[10] M. Li, J. Wu, D. Wang, T. Zhou, Z. Di, and Y. Fan, "Evolving model of weighted networks inspired by scientific collaboration networks," *arXiv:cond-mat*, paper number 0501655, 2005.
[11] W. Aiello, F. Chung, and L. Lu, "A random graph model for massive graphs," *Proc. 32nd Annual ACM Symposium on Theory of Computing*, ACM, New York, USA, 2000, pp. 171–180.
[12] W. Wang, B. Wang, B. Hu, G. Yan, and Q. Qu, "General dynamics of topology and traffic on weighted technological networks," *Phys. Rev. Lett.*, vol. 94, 188702-1–4, May 2005.
[13] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, January 2002.