

动态短信通信复杂网络演化模型研究

孙钦东^{1,2}, 孙亚红¹, 管晓宏², 王倩¹

(1. 西安理工大学计算机科学与工程学院, 710048, 西安; 2. 西安交通大学
智能网络与网络安全教育部重点实验室, 710049, 西安)

摘要: 为了深入理解手机短信通信网络的演化规律,通过分析手机短信通信网络生长过程的特点,研究了已有复杂网络演化模型刻画短信通信网络生长过程的不足,从而建立了短信通信复杂网络的动态演化模型. 该模型在考虑网络动态增长的同时,兼顾了网络内部边的动态演化以及节点的退出,更细致、真实地刻画了短信网络的动态演化特性. 实验结果显示,该模型所生成的网络规模表现出有增有减的动态演化过程,与现实情况相吻合,能够更加客观地模拟与刻画短信复杂网络的特征与演化状况.

关键词: 短信; 复杂网络; 演化模型

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2009)06-0005-05

A Dynamic Evolution Model of Short Message Complex Network

SUN Qindong^{1,2}, SUN Yahong¹, GUAN Xiaohong², WANG Qian¹

(1. School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China, 2. MOE Key Laboratory for Intelligent and Network Security, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The law of SMS (short message service) network growth is analyzed to understand the evolution of SMS communication network, and the deficiencies of existing complex network models that are used to describe real SMS network are pointed out. Then, a new dynamic evolution model is proposed. The model considers both the network node removing and the inner and outer growth of the network, so that the dynamic evolution of a true network can be described in more detail. Experimental results show that the proposed model describes the growth of SMS network desirably, and the model can more objectively simulate and depict the features and evolution of the SMS complex network.

Keywords: short message; complex network; evolution model

已有研究表明,在自然界与社会生活中,众多的复杂系统可以用复杂网络加以描述^[1],并在社会关系网络、互联网拓扑结构、生物神经网络系统及国际贸易往来关系等真实网络中得到了广泛应用^[2-4],因为它能揭示大量的真实系统表现出的小世界效应、无尺度特性以及聚集性等.

目前,手机短信已成为继 Internet 之后的“第五媒体”,成为人们日常交流的主要方式之一^[5]. 短信传播网络兼具了大众传播与人际传播的特点,为大

规模突发的热点事件传播提供了更为便捷的平台,同时也形成了一种新型的复杂网络模式. 如何从本质上揭示短信通信复杂网络的发展与演化规律,研究短信通信网络的演化机制以及网络的拓扑特性,是实现短信复杂网络管理与控制的基础,也是短信舆情分析及预警技术中的关键性基础理论问题.

少数研究者在这方面开展了有益的研究. 文献[6-7]分别用小世界模型^[8]与 Barabasi-Albert (BA)模型^[9]来研究短信的传播,揭示了短信传播过程的

一些特性.但是,小世界模型和BA模型是复杂网络研究中最基本的模型,用它们来刻画短信通信网络,不能完全揭示短信通信网络所独有的特点.文献[10]基于真实的短信通信数据,在对节点度进行详细分析的基础上,提出了基于局部优先连接机制的短信网络生长模型,而文献[11]通过分析不良短信的传播特点,建立了谣言短信传播网络模型.这2篇文献均从网络增长的局部角度出发,来建立网络的生长模型,但对于刻画网络宏观生长过程中的一些特性略显不足.文献[12]根据真实的电话呼叫关系建立了一种加权网络,研究了网络中边强度的变化对网络结构的影响,而文献[13]也是一种针对电话呼叫网络的实证性展开的研究,分析了网络的子图连通性、膨胀率、度分布等结构特性.因为电话呼叫网络与短信通信网络既有联系,又有区别,所以这2篇文献为本文的研究提供了有意义的参考.可见,现有研究均从特定的角度对短信网络进行了分析,其研究成果在一定程度上反映了短信传播过程的特点,但对于刻画真实的短信网络有一定的局限性.

本文通过分析现有网络模型的构造原理,分析了上述模型在刻画短信网络上的不足,进而建立了一种能充分考虑短信网络内部演化和节点退出现象的网络模型,以期从宏观角度来揭示短信网络的拓扑特性和短信网络的演化规律.

1 已有短信复杂网络模型分析

BA模型^[9]的提出具有重大的意义,它首次把幂律度分布引入到了网络,从演化的角度研究了网络的一些宏观性质的起源,为复杂网络的研究提供了一种新的视角和思路,在一定程度上刻画了一些现实网络.但是,BA模型只能生成度分布指数为3的网络^[11],而在实证研究中发现,短信网络的度分布指数小于3,因此在刻画短信通信网络上有一定的局限性.

文献[10]通过对短信网络真实数据的分析,建立了基于局部优先连接的短信网络生长模型,从其分析可以看出,模型在度分布方面能很好地刻画短信网络的动态演化特性,但是忽略了这样一种情况:网络中的某一个节点,不仅以一定的概率认识新加入网络的节点,更多的是它和已存在于网络中的节点建立了关联关系,即网络内部演化,因此用它刻画短信网络的动态演化也有一定的局限性.

综合上述分析可以看出,现有的模型在刻画复杂网络时,网络的规模是随着时间推移逐步递增的,

即在网络规模增长的同时都会有一个节点加入到网络中来,这与现实情况是不相符的.在现实情况中总有这样一种情况:手机用户可能因为网络服务、资费等问题而退出当前网络,在短信通信网络中表现为节点退出网络,移动通信运营的数据也验证了这个问题.

因此,针对现有模型的缺陷,本文提出了兼具内部演化和节点退出的短信网络演化模型,以更好地刻画实际短信通信网络的演变.

2 兼具内部演化和节点退出的短信网络演化模型

根据实际的短信通信网络交互过程,以手机号码为节点、以用户之间的短信发送关系为边,建立短信通信网络.从直观上看,短信通信网络的规模不是简单连续增长的,在增长过程中,不仅有新节点、新连线加入,也会有旧节点、旧连线被删除,表现出有增有减的动态演化.

在移动通信业务发展初期,只有少数用户之间存在短信往来,这是短信网络的初始状态.随着短信业务的推广,越来越多的人使用短信进行交流,即越来越多的节点加入短信通信网络,它促使了短信通信网络的发展.

从短信通信网络内部来看,网络中每一个用户(节点)不仅可以认识新用户,也可能因为某种原因认识网络中朋友的朋友,从而在网络内用户之间建立了短信通信关系,由此导致了短信网络的内部演化.另一方面,短信通信网络是建立在网络运营服务之上的,运营商之间存在着竞争,用户可能会因此变换服务商,如用户因资费问题更换网络运营商等,在短信通信网络中则表现为用户退出网络,这使得短信网络中有节点退出.

综上,实际的短信通信网络应该是一种兼具内外节点增加和退出的复杂网络,就此本文建立如下的网络生长模型.

(1)初始时刻,网络中有 n_0 个节点、 m_0 条边,节点之间以任意方式相连.

(2)在每一个时间步,产生一个 $(0,1)$ 之间均匀分布的随机数 p .

(3)若 $p < p_0$ (p_0 随机指定),从网络中随机选择一个节点,并将其去除,同时与该节点相关联的边也被去除.

(4)若 $p \geq p_0$,先在网络内部以概率 q 增加 n 条

新边,新增边的一端随机选取,另一端按照节点度优先连接进行选取,并重复 n 次;将一个新增节点加入到网络中,并与 $m(m < n_0)$ 个已经存在的节点建立连接. 新增节点与网络中已存节点 i 以概率 Π_i 进行连接, Π_i 的表达式为

$$\Pi_i = \frac{k_i}{\sum_j k_j} \quad (1)$$

式中: k_i 为节点 i 的度; k_j 为与节点 i 相连的节点 j 的度. 可以看出,本文模型在考虑网络规模动态增长过程的同时,兼顾了网络内部边的动态演化,较之于已有的短信通信网络模型,能更加细致、真实地刻画网络的动态演化特性.

3 模型生成的网络特征分析

3.1 网络度分布特性分析

使用 MATLAB 平台对所提模型的网络度 k 进行了数值模拟,结果如图 1~图 3 所示. 可以看出,随着 p_0 、 q 的变化,网络度分布状况发生了一系列的变化.

当 $p_0=0$ 、 $q=0$ 时,本文模型为 BA 模型;当 $p_0=0$ 、 $q \neq 0$ 、网络规模线性增长时,本文模型为一种兼顾网络内部动态演化的网络模型. q 由 0 变化至 1,网络的累积度 $P(k)$ 的变化如图 1 所示.

由图 1 可以看出,当没有节点退出网络时,网络规模的增长是绝对的,此时只有节点数和内部边在增加. 随着 q 的逐渐增大,网络度分布曲线的变化越来越平缓,即幂指数逐渐减小. 这说明当一个人以较大的概率认识其他朋友的时候,整个网络里朋友多的人也就多了,亦即网络中节点的度普遍增大,在现实中也是如此.

当 $p_0 \neq 0$ 、 $q=0$ 时,本文模型的网络度分布曲线随着 p_0 的增加而变化的曲线如图 2 所示.

由图 2 可以看出,当 $q=0$ 、 $0 \leq p_0 \leq 0.1$ 时,亦即当以概率为 0.1 删除网络中的部分节点时,网络可以保持完好的连通性,且网络度分布表现为幂律形式. 这说明在网络的演化过程中,若以较小的概率随机移除网络节点时,不会影响网络的幂律分布. 当 $p_0 > 0.1$ 时,网络度分布不再表现为幂律形式,度曲线在其值大于 10 的地方发生了严重偏离. 当 $p_0 > 0.25$ 时,在多次实验中均发生了网络不连通现象,且随着 p_0 的增加,网络不连通的概率越来越大. 这说明当以较小的概率随机去除网络中的部分节点时,不会影响网络的连通性,体现出网络具有对随机攻击的鲁棒性.

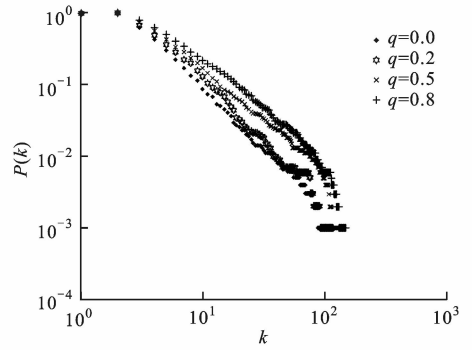


图 1 $p_0=0$ 、 q 逐渐增大时网络累积度分布状况

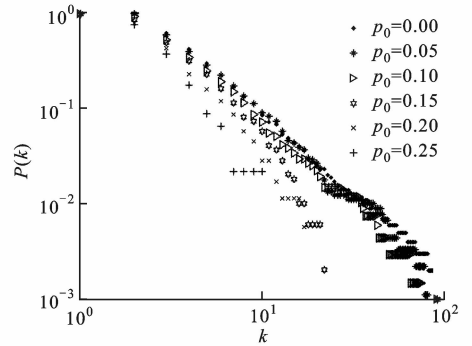


图 2 $q=0$ 、 p_0 逐渐增大时网络累积度分布状况

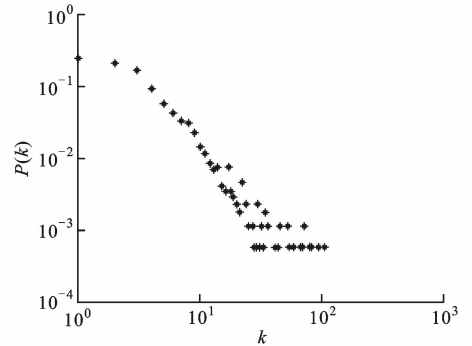


图 3 $p_0=0.02$ 、 $q=0.5$ 、 $n=2$ 时网络累积度在双对数坐标下的分布状况

为了保证网络的连通性,基于上面的分析,初次断定 p_0 的取值范围为 $0 \sim 0.25$. 由于随着 q 的增加,网络度分布指数不断减小,因此在本模型中 q 取值应尽可能大. 通过不断调节 p_0 、 q 、 n ,最终得到 $p_0=0.02$ 、 $q=0.5$ 、 $n=2$ 时,网络度分布的幂律指数为 2.76,这样度分布曲线在其值较小的地方才会发生偏离,如图 3 所示. 该结果与文献[5]的实证研究结果相吻合.

3.2 网络的平均路径长度

在网络研究中,2个节点间的距离定义为连接2节点最短路径边的数目.网络的直径为任意2节点间的最大距离,网络的平均路径长度 L 则为所有节点对之间的距离平均值,它可描述网络中节点间的分离程度.当取模型参数 $p_0=0.02$ 、 $q=0.5$ 、 $n=2$ 时,可生成网络节点数分别为1 000、2 000、3 000的3个网络,通过计算得到的网络平均路径长度分别为4.83、5.64、6.23.也就是说,在这3个网络中,连接任意2节点之间的最短路径的平均边数分别为4.83、5.64、6.23,可以看出该模型生成网络的平均路径是非常小的.

3.3 网络的聚集系数

聚集系数 C 可以描述网络中节点的聚集情况.假设节点 i 通过边与其他 r_i 个节点连接,如果这 r_i 个节点均相互连接,它们之间应存在 $r_i(r_i-1)/2$ 条边,而这 r_i 个节点之间存在的边数为 E_i ,则 E_i 与 r_i

$(r_i-1)/2$ 之比就是节点 i 的聚集系数 C_i .整个网络中所有节点的聚集系数的平均值就称为网络的聚集系数 C .利用本文模型生成网路,其聚集系数分别为0.32、0.43、0.49,可以看出该短信网络聚类性很好.

4 与现有模型的比较及分析

在MATLAB中,把现有的BA模型、局部优先连接模型及本模型通过编程实现,并生成节点数(网络规模)为1 000、2 000、3 000的网络,然后在Pajek中计算网络的各相关特性参数.在生成文献[10]的局部优先连接模型时,根据文中结论知, p_0 的取值不影响网络度的分布,因此网络的初始状态随机给定, p_1 与 p_2 这2个参数分别为0.2、0.06.在生成BA模型时,网络的初始状态、参数 m 的取值不影响网络度的分布,因此在保证 $m \leq m_0$ 的前提下可随意给定.根据第3节中的实验分析,本模型的 p_0 、 q 、 n 分别为0.02、0.5、2.网络特性的对比结果如表1、图4所示.

表1 3种模型生成的不同网络规模(节点数)的网络特性比较

BA模型				局部优先连接模型				本文模型				节点数
L	C_i	D	$\bar{k}$	L	C_i	D	$\bar{k}$	L	C_i	D	$\bar{k}$	
5.98	0.052	3	6.01	2.64	0.537	2.78	4.22	4.83	0.32	2.73	3.74	1 000
6.27	0.057	3	5.99	3.33	0.610	2.73	4.31	5.64	0.43	2.74	3.94	2 000
6.35	0.061	3	6.10	4.17	0.594	2.76	4.43	6.23	0.49	2.76	4.12	3 000

注: D 、 $\bar{k}$ 分别表示度分布指数、平均度.

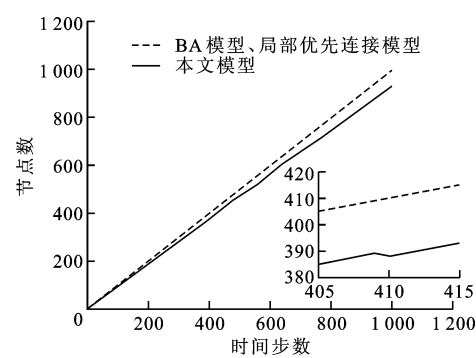


图4 网络规模随时间步的变化

由以上数据可以看出,本模型与局部优先连接模型生成的网络度分布指数均小于BA模型,其范围在2~3之间,这与实际网络是相符的.在平均路径长度、聚集系数、度分布指数、平均度等网络特性的刻画上,本模型与局部优先连接模型相当.从网络规模增长过程来看,局部优先连接模型生成的网络

度是递增的,而本文模型生成的网络度不再是递增的,表现出了有增有减的动态变化过程.从网络规模的变化上看,BA模型和局部优先连接模型生成的网络规模都是直线递增的,本模型生成的网络规模表现出了有增有减的动态演化过程(如图4所示),这与现实情况相吻合.因此,用本文模型来仿真短信通信网络更为合理.

5 结 论

本文通过对已有BA模型和局部优先连模型进行分析,指出了它们在刻画手机短信网络动态演化特性的局限性,分析了短信网络生长过程的特点,建立了综合考虑网络节点增长、内部演化以及节点退出等问题的动态演化模型.通过调节参数,得到了与已实证的研究结果相符的参数,并据此生成不同规模的短信网络.与现有模型的聚集系数、平均路径长度等复杂网络特性参数进行比较,结果表明本文模

型能够很好地刻画短信网络的演化特性和演化规律。

今后的工作是,进一步研究节点属性(即手机用户的活跃程度、知识状态等)、关系方向(短信发送、接收方向)、关系权重(一定时刻短信发送的数量)等,建立一种加权有向网络,以期更好地揭示短信通信网络的本质特性。

参考文献:

- [1] COSTA L F, RODRIGUES F A, TRAVIESO G, et al. Characterization of complex networks: a survey of measurements [J]. *Advances in Physics*, 2007, 56(1): 167-242.
- [2] BOCCALETTI S, LATORA V, MORENO Y, et al. Complex networks: structure and dynamics [J]. *Physics Reports*, 2006, 424: 175-308.
- [3] MOTTER A E, MÁTIAS M A, KURTHS J, et al. Dynamics on complex networks and application [J]. *Physica: D*, 2006, 224(1/2): 7-8.
- [4] LI Xiang, CHEN Guanrong. A local world evolving network model [J]. *Physica: A*, 2003, 328(1/2): 274-286.
- [5] 刘思林. 手机不良短信的社会影响及其综合治理策略 [J]. *北京邮电大学学报(社会科学版)*, 2004, 6(1): 53-56.
LIU Silin. On harmful effects of short messages of mobile telephone on the society and their treatment in a comprehensive way [J]. *Journal of BUPT(Social Sciences Edition)*, 2004, 6(1): 53-56.
- [6] 王静,孔令江,刘慕仁. 小世界网络上的手机短信息传播模型 [J]. *广西师范大学学报(自然科学版)*, 2006, 24(3): 2-4.
WANG Jing, KONG Lingjiang, LIU Muren. Model for spread of mobile-telephone short message on small world networks [J]. *Journal of Guangxi Normal University (Natural Science Edition)*, 2006, 24(3): 2-4.
- [7] 王静,孔令江,刘慕仁. BA 网络上的手机短信息传播模型 [J]. *广西物理*, 2006, 27(1): 20-21.
WANG Jing, KONG Lingjiang, LIU Muren. Model for spread of mobile-telephone short message on BA networks [J]. *Physics of Guangxi*, 2006, 27(1): 20-21.
- [8] WANG Xiaofan, CHEN Guanrong. Complex networks: small world, scale free and beyond [J]. *IEEE Circuits and System Magazine*, 2003, 3(1): 6-21.
- [9] HAYASHI Y. A review of recent studies of geographical scale-free networks [J]. *IPSJ Digital Courier*, 2006, 1(2): 155-164.
- [10] 吴晔,肖井华,马宝军,等. 手机短信网络的生长过程研究 [J]. *物理学报*, 2007, 56(4): 2037-2041.
WU Ye, XIAO Jinghua, MA Baojun, et al. Research on the growing process of short message networks [J]. *Acta Phys Sin*, 2007, 56(4): 2037-2041.
- [11] 徐增勇,宋运忠. 谣言短信传播网络拓扑性质研究 [J]. *河南理工大学学报(自然科学版)*, 2008, 27(2): 217-223.
XU Zengyong, SONG Yunzhong. On topological properties of propagation networks for rumor shot message [J]. *Journal of Henan Polytechnic University (Natural Science Edition)*, 2008, 27(2): 217-223.
- [12] ONNELA J P, SARAMAKI J, HYVONEN J, et al. Structure and tie strengths in mobile communication networks [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(18): 7332-7336.
- [13] 王艳辉,吴斌,王柏. 电信社群网络静态几何性质分析 [J]. *复杂系统与复杂性科学*, 2005, 2(2): 54-60.
WANG Yanhui, WU Bin, WANG Bai. Research on static measures of telecom society networks [J]. *Complex Systems and Complexity Science*, 2005, 2(2): 54-60.

(编辑 苗凌)

【本刊相关文献链接】

- 一种可演化网络的研究与实践. 2008, 42(10): 1193-1203.
- 基于复杂网络理论的流程工业系统安全性分析. 2007, 41(7): 806-810.
- 基于先验知识与模块性的网络社区结构探测算法. 2007, 41(6): 750-754.
- 基于小世界网络的复杂系统故障传播模型. 2007, 41(3): 334-338.
- 基于复杂加权网络的移动协同空间分析. 2006, 40(5): 573-576.
- 一种基于蚁群聚类的径向基神经网络. 2006, 40(4): 386-389.