基于拷贝模型的复杂网络鲁棒性研究 *

黄金源[1], 张 宁[1], 肖仰华[2]
(1. 上海理工大学 管理学院,上海 200093; 2. 复旦大学 计算机科学技术学院,上海 200433)

摘 要: 拷贝机制被广泛认为是系统通过构造冗余提高自身鲁棒性的主要机制之一。为了探究拷贝机制是否是真实网络鲁棒的基本机制,通过计算机仿真的方法对基于拷贝机制的网络模型的鲁棒性进行了进一步的研究。仿真结果表明,随着拷贝机制的增强,相应网络对于随机故障的鲁棒性增强,而对于蓄意攻击的鲁棒性减弱。这一事实启发人们,基于拷贝机制的网络,其鲁棒性仅局限于随机失效,单纯的拷贝机制还不足以有效抵抗基于网络全局结构信息的蓄意攻击。

关键词: 复杂网络; 鲁棒性; 拷贝机制; 结构冗余
中图分类号: TP393 **文献标志码**: A **文章编号**: 1001-3695(2010)04-1403-04
doi:10.3969/j.issn.1001-3695.2010.04.054

Study of robustness of complex networks based on copying model

HUANG Jin-yuan[1], ZHANG Ning[1], XIAO Yang-hua[2]
(1. *School of Management*, *University of Shanghai for Science & Technology*, *Shanghai* 200093, *China*; 2. *School of Computer Science*, *Fudan University*, *Shanghai* 200433, *China*)

Abstract: Copy mechanism is widely believed to be one of the main mechanisms by which system can improve its robustness. To investigate whether the copy mechanism is the basic mechanism accounting for the robustness of real networks, this paper studied the robustness of networks generated by copying model by computer simulation. The simulation result shows that the stronger the copy mechanism is, the more robustness the corresponding network is under random failure, but the more vulnerable to intentional attack. Such facts imply that robustness of copying network is limited to its robustness under random failure, and that pure copying mechanism is not enough for a network to be robust against intentional attack based upon global structural information of the network.

Key words: complex networks; robustness; copy mechanism; structural redundancy

0 引言

自小世界网络模型[1]以及无标度网络模型[2]问世以来,面向各类真实网络的统计性质的研究以及对网络功能与行为的研究逐渐成为包括数学、物理学、生物学以及信息科学等各个学科领域的研究热点,并逐渐形成独立的学科分支,即所谓的网络科学。

经过十来年的发展,网络科学已经形成很多具体的研究方向,如网络模型[3]、社区结构[4]、网络同步[5]、网络谱分析[6]、网络对称性[7]等。其中一个很重要的研究问题是网络的鲁棒性分析[8]。在一个典型的网络鲁棒性分析问题中,人们一般比较关心在受到来自环境的随机干扰以及外界的恶意攻击时,网络还能否保持其必要的结构性质不变,从而依旧维持原有功能。通常,如果网络组件,也就是点或边,在遭受随机故障或蓄意攻击时依旧能保持网络的连通性,或者说网络组件的失效对信息传输能力影响较小,那么这个网络则被视为鲁棒的。随机故障等价于在没有任何网络结构信息的前提下对网络展开随机攻击。而一旦掌握一定的网络结构信息,攻击者则会采取破坏力更强的蓄意攻击的方式,即优先破坏网络中重要的节点或边。

目前学术界已经针对网络鲁棒性开展了一定程度的研究,其主要手段包括仿真分析和解析分析。其中,Albert 等人[8]较早针对随机网络[9]、无标度网络及真实网络(国际互联网、万维网)的鲁棒性展开了实证对比研究,发现无标度网络具有随机故障下鲁棒、蓄意攻击下脆弱的双重特性,而这种性质从根本上讲取决于无标度网络度分布的异构性。之后,鲁棒性的研究一直是网络研究热点之一,其中 Holme 等人[10]系统地研究了在不同的攻击策略下,多种网络功能指标的变化所体现出的网络鲁棒性。另外,Cohen 等人[11]利用广义随机图上的渗流理论,给出了特定度分布的网络在基于度的蓄意攻击和随机失效下完全不连通时,所需要移出的节点百分比的阈值。Paul[12~15]等人最近的一些工作考虑了如何以最小的代价对网络的鲁棒性进行优化。而国内,张宁等人[16~18]研究了不同条件下计算机病毒对复杂网络鲁棒性的影响,并针对现实世界网络传播的特性,对不同网络结构的计算机病毒的传播和控制策略进行了探讨,提出了在资源有限的条件下复杂网络控制病毒

收稿日期: 2009-08-25; **修回日期**: 2009-09-26 **基金项目**: 国家自然科学基金资助项目(70971089);上海市重点学科建设资助项目(S30501);上海市研究生创新基金资助项目(JWCXSL0902);河南省教育厅自然科学研究计划资助项目(2009A520023)

作者简介: 黄金源(1985-),男,山东泰安人,硕士研究生,主要研究方向为复杂网络(jinyuan.h@gmail.com);张宁(1956-),女,江苏人,副教授,硕士,主要研究方向为复杂网络;肖仰华(1980-),男,江苏人,讲师,博士,主要研究方向为图数据库、复杂网络.

传播的有效策略。谭跃进等人[19,20]研究了网络拓扑结构对系统抗毁性的影响,并提出了基于多 agent 建模仿真的复杂网络鲁棒性研究方法。

然而,一直以来,一个与鲁棒性密切相关的网络模型—拷贝模型[21]的研究却尚未见报道。拷贝或复制被广泛认为是生物系统以及相应的各类生物网络借以提高自身鲁棒性的主要机制[22]。这一机制具体到网络模型中,则体现为拷贝模型。那么,人们自然会直观地认为,拷贝模型会导致网络较高的鲁棒性。然而,这一经验判断一直未得到定性或定量证实。本文的工作即致力于验证拷贝模型的鲁棒性。通过理论分析和大量的实验,笔者发现与经验知识相反,拷贝模型在基于度和介数的蓄意攻击下,在一系列网络关键功能指标上,体现出较为明显的脆弱性,其仅在随机攻击下在某些网络功能指标上体现出一定的鲁棒性。

1 拷贝模型

拷贝模型最早来自计算机科学对于万维网图结构模拟的研究。万维网图可以看做一个以静态网页为节点,以一个网页到另一个网页的超链接为边的网络。当一个新的静态页面加入到万维网图中,该页面倾向于链接到内容主题与之相近的页面所链接的页面。也就是说,新页面会有一定倾向性地选择万维网图中的某个页面作为原型节点,并拷贝其部分或者全部链接。这些事实构成了拷贝模型的基本思想。值得注意的是,拷贝机制不仅存在于万维网中,它也是一种在生物学领域常见的生物演化机制,在生物网络中点或者边的拷贝现象也是普遍存在的,并被认为是生物系统改善其抗环境干扰或者提高自身鲁棒性的主要机制之一,如前面曾提到的复制模型[22]。本文即以拷贝模型为代表研究点边的拷贝机制对于网络鲁棒性的影响。

一个简单的拷贝模型需要两个参数,即拷贝因子 α 和节点出度常数 $d(0<\alpha<1, d\geqslant 1)$。其中:$\alpha$ 用于控制拷贝原型节点链接对象的程度,d 用于控制网络的浓密程度。在每一个时间步,将一个新节点 u 加入到现有网络中,并指定 u 连接 d 个现有节点。首先要为 u 在现有节点 $V(t)$ 中随机选择一个原型节点,从 u 连出的第 $i(0<i\leqslant d)$ 条边则以 α 的概率随机地从 $V(t)$ 中选择连接,以 $1-\alpha$ 的概率与原型节点的第 i 个邻接点相连接。这样的过程持续下去直到达到指定的网络规模为止。

由于拷贝模型考虑了选取原型节点的邻接点作为新节点连接的对象,从而容易导致新节点与原型节点共同连接部分相同的邻接点,而邻接点相同或相似的节点通常被认为是结构相似的[7]。那么人们会自然地认为拷贝行为容易产生更多的网络结构冗余,而结构冗余直觉上会带来网络鲁棒性。如图 1 所示,10 号点以 6 号点为原型,拷贝了其全部连接目标,从而使得 6 号点与 10 号点有着完全相同的邻接点,那么这两个节点在常见节点度量下都有着相同的取值。因而在此种意义下,这两个节点可以被视做是结构等价的。事实上,在图 1 中,6 号点与 10 号点是自映射等价的,是一种在最严格意义下的节点等价关系[7],而结构等价通常蕴涵着功能等价[23]。例如,假设图 1 展示的是一个社会网络,则 6 号点与 10 号点所代表的实体很可能在整个网络系统中扮演相同或相似的社会地位。因此,它们在结构上或功能是冗余的,即可相互替代。那么当这两个点中的一个节点受到攻击,另一个结构等价的节点将履行被攻击节点的功能,从而有效保持这些节点所处的局域结构的功能。在生物网络中也存在类似的情形。正是基于上述分析,人们通常认为,拷贝导致冗余,冗余带来健壮。

2 攻击方式和功能指标

上述关于拷贝模型鲁棒性的分析忽视了攻击者攻击网络的方式以及网络组件失效的方式。在理论上,如果攻击者每次攻击的对象都是结构上有备份的节点,那么基于拷贝模型的网络预计将具有较好的抗攻击能力。但事实上,攻击者的目标是以最小代价取得最大破坏效果,因此不可能采取这样一种不利于自己的攻击方式。一方面,如果攻击者能够获得整个网络的结构等价信息,攻击者可以攻击不存在结构上与之等价的冗余对象的节点,此时网络很可能受到重大破坏。另一方面,如果攻击者不能够获得整个网络的结构等价信息,那么攻击者通常会采取信息获取代价相对较小的基于节点度或介数的蓄意攻击或者无须任何信息的随机攻击,在这两种攻击方式下,拷贝网络是否还能体现出较为明显的鲁棒性,仍是尚未解决的问题。本文后面的内容将回答这一问题。

把握攻击方式和网络功能指标的演化行为是理解网络鲁棒性的关键。在随机故障中,攻击者将采取完全随机的方式删除网络中的节点。在蓄意攻击中,攻击者会根据所得信息将节点按照重要性从大到小排序并以此顺序进行攻击,以期以较小代价获得较大破坏性。节点的度与介数是表征网络中节点重要程度的两个常用测度。在网络中,度大的节点对于保持整个网络的连通性至关重要,而介数较大的节点是保持网络信息传递功能的重要节点。

网络的功能指标,通常包括最短路径相关指标,即平均最短路径 P 和网络直径 D。当整个网络受到攻击时其内部有可能不再连通,因此本文的平均最短路径采用文献[10]中的定义:

$$L^{-1}=\frac{1}{N(N-1)}\sum_{i\in V,j\neq i\in V}\sum d_{ij}^{-1}$$

其中:V 为网络的点集;$N=|V|$;d_{ij} 为点 i、j 之间的最短长度。令 $P=L$,即 P 为网络的平均最短路径。显然,当节点受到攻击后,P 与 D 会增大,其增大幅度越小,网络越健壮。

网络功能另一项关键指标是连通分量相关指标,其中最主要的是最大连通分量的规模,记为 $S_{\max}$。显然,某个节点失效后,$S_{\max}$ 变化越小网络越健壮。此外,试验中还将用到平均连通分量的规模,记为 S_{avg},及连通分量的数目,记为 M。

3 鲁棒性的模拟分析

为了测试拷贝模型的鲁棒性,以 10 个节点的完全图作为初始网络,令节点出度常数 $d=3$,时间步 $t=2\ 990$(连续添加 2 990个节点),并分别令 $\alpha=0.3$、0.6、0.9,从而得到三个规模相同($N(\text{node})=3\ 000$,$E(\text{edge})=9\ 015$),却有着不同程度结构冗余的网络数据,分别记为 $G_1(\alpha=0.3)$,$G_2(\alpha=0.6)$,$G_3(\alpha=$

0.9)。根据前文所述的拷贝模型生成机制,随着拷贝因子 α 的递增,G_1、G_2、G_3 中的结构冗余规模将会递减。若令 $SR(G)$ 代表网络 G 中结构冗余规模的大小,即 $SR(G)$ 越大,网络 G 的结构冗余的规模越大,则有 $SR(G_1) > SR(G_2) > SR(G_3)$。接下来将对结构冗余规模依次变化的三个网络进行不同的攻击仿真实验,从而考察拷贝模型导致的结构冗余对于网络鲁棒性的影响。

3.1　蓄意攻击

在模拟蓄意攻击时,分别采用基于节点度和介数的两种重要程度参数来选择优先攻击的节点。对三个网络分别按照节点度值和介数由大到小的顺序进行去点操作,每次去点后计算该网络的网络功能指标,并重新对网络中剩余节点再次按照两个测度的大小进行排序,再次去点。通过对网络反复进行持续攻击操作发现,三者在各个测度上表现出明显不同的趋势。

在度或介数较大的节点受到持续攻击时,网络中某两个节点之间的最短路径长度会由于其邻接点的受损而逐渐增长,因而整个网络的最短路径会随之增大[8]。图 2、3 中的三个不同拷贝参数下的网络都符合这一事实。但是,三个网络在最短路径的增长速度上体现出截然不同的趋势。显然,G_1 增长最快,G_2 与 G_3 依次渐缓。这说明,随着拷贝因子 α 的减小,网络体现于最短路径方面的功能越容易受到破坏。在网络直径方面,拷贝因子对于网络鲁棒性的作用与最短路径长度相似。如图 4,三个网络在受到基于介数的蓄意攻击时,G_1 的网络直径增长速度明显快于 G_2 和 G_3,G_3 的网络直径增长速度最慢。上述两个对于网络路径方面的统计指标与拷贝因子之间的关系充分说明了在网络结构形成过程中,拷贝的程度越大,最终生成的网络受到上述蓄意攻击时在维持通信功能的能力上越弱。

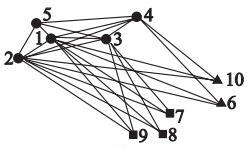

图1 拷贝模型产生结构冗余节点的示意图

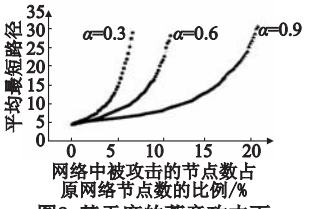

图2 基于度的蓄意攻击下平均最短路径的变化情况

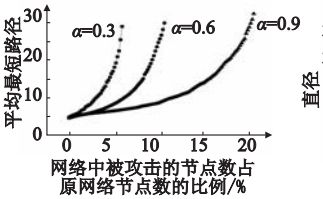

图3 基于介数的蓄意攻击下平均最短路径的变化情况

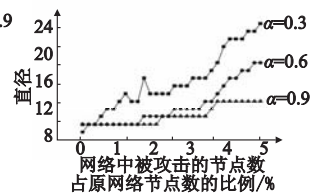

图4 基于介数的蓄意攻击下直径的变化情况

接下来,进一步考察拷贝因子与网络连通性相关指标之间的关系。

一般而言,当网络中关键节点受损时,其最大连通分量 S_{max} 会逐渐变小,而网络中连通分量的数目 M 却会逐渐增多[8],即一个网络在遭受蓄意攻击时,其最大的连通子图会渐渐崩溃,与此同时会出现更多较小的子图,整个网络变得支离破碎。如图 5 ~ 8 所示,在遭受蓄意攻击时,三个网络的 S_{max} 均逐渐减小,而 M 均逐渐增多。但是,显然不同拷贝因子下的网络在受到攻击时,其相应 S_{max} 的衰减行为完全不同:拷贝因子越小,S_{max} 衰减到 0 的速度越快,且致使 S_{max} 最终下降为 0 所需攻击的网络节点的比例最小。也就是说,对于 G_1,使得整个网络所有节点彼此之间不再相互连通的攻击代价是最小的。相反,对于 G_2 和 G_3,使得整个网络支离破碎所需的攻击代价则依次更大。通过观察三个网络 M 的变化情况也可获知,G_1 在遭受蓄意攻击时比 G_2、G_3 更容易变得支离破碎。

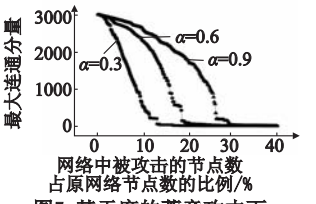

图5 基于度的蓄意攻击下最大连通分量的变化情况

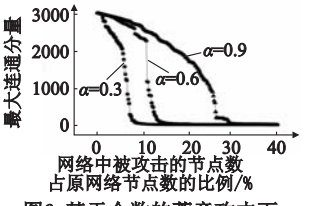

图6 基于介数的蓄意攻击下最大连通分量的变化情况

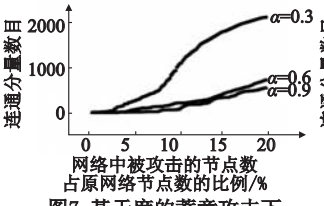

图7 基于度的蓄意攻击下连通分量数目的变化情况

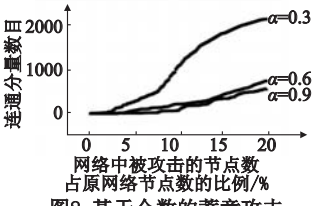

图8 基于介数的蓄意攻击下连通分量数目的变化情况

以上实验现象说明,在遭受蓄意攻击时,结构冗余规模最大的网络 G_1 的鲁棒性最差,而 G_3 表现出的鲁棒性最好。如果将网络在蓄意攻击下的鲁棒性程度简记为 R_a,则有

$$R_a(G_1) < R_a(G_2) < R_a(G_3)$$

3.2　随机故障

在本节,将进一步研究拷贝网络在随机故障模式下网络的鲁棒性表现。

文献[8]已证实,对于无标度网络,在随机故障模式下其平均最短路径和直径的变化趋势与基于度的蓄意攻击相似,会随着故障节点的增多而变大,但是其增长速度相对缓慢。由于本文使用的拷贝网络的度分布也满足幂律分布[21],因而三个网络在遭受连续随机故障时,P 和 D 都是逐渐增大的,如图 9、10 所示。但是,G_1 的两个测度的增长速度最慢,而 G_3 最快。也就是说拷贝因子越小的网络在遭受随机故障时,其网络最短路径相关指标被破坏程度相对较小。

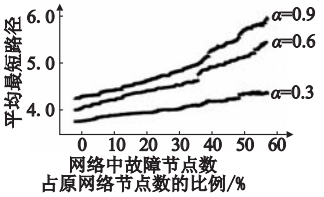

图9　随机故障下平均最短路径的变化情况

为了进一步验证随机故障模式下拷贝因子与网络鲁棒性的关系,笔者还考察了平均连通分量的大小 S_{avg} 在遭受随机故障时的变化趋势。这里的平均连通分量采用文献[8]中的定义。一般而言,当网络受到蓄意攻击或随机故障时,某个节点的移出将会导致越来越多的子网络脱离最大连通分量,因而平均连通分量在网络结构迅速瓦解的阶段会表现出增大的趋势[8]。如图 11 所示,三个网络的 S_{avg} 在连续的随机故障下均逐渐增大。但是,仍然观察到不同拷贝因子的网络呈现出的完全不同的变化趋势,显然 G_1 的增长速度慢于后两者,而 G_3 的增

长速度最快。这说明,拷贝因子越小的网络在随机故障下呈现出一定的鲁棒性,与上文的结论相吻合。

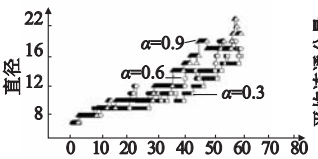

图10 随机故障下直径的变化情况

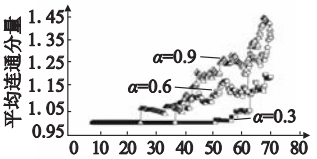

图11 随机故障下平均连通分量的变化情况

以上实验现象表明,在遭受随机故障时,结构冗余规模最大的网络 G_1 的鲁棒性最好,而 G_3 表现出的鲁棒性最差。如果将网络在随机故障下的鲁棒性程度简记为 R_f,则有

$$R_f(G_1) > R_f(G_2) > R_f(G_3)$$

3.3 结果分析

上述实验说明,不同拷贝因子下的拷贝网络在各种攻击模式以及网络功能度量指标下均呈现出明显不同的鲁棒性行为。这一事实充分说明拷贝机制对于网络鲁棒性有着重要的不可忽视的影响。如表 1 所示,随着拷贝因子的减小、网络规模保持不变的前提下,网络的结构冗余不断增加,正是这种结构冗余导致了网络的不同鲁棒性特征。

表 1 鲁棒性实验结果比较

实验项	结果
拷贝因子 α	$0.3 < 0.6 < 0.9$
拷贝因子 α 所致结构差异	$SR(G_1) > SR(G_2) > SR(G_3)$
蓄意攻击下的鲁棒性	$R_a(G_1) < R_a(G_2) < R_a(G_3)$
随机故障下的鲁棒性	$R_f(G_1) > R_f(G_2) > R_f(G_3)$

拷贝模型在随机故障和蓄意攻击下截然相反的鲁棒性特性对人们有一定的启发,即结构冗余所带来的鲁棒性仅限于基于有限网络结构信息的攻击,如随机故障或局部攻击。而对于基于全局网络结构信息的攻击手段,如基于度或者介数的持续蓄意攻击,拷贝网络则显得非常脆弱。这一事实与人们的经验知识恰好相反,但却可以得到合理解释。注意到在网络规模不变的前提下,提高网络的结构相似性,将直接导致网络的度分布越来越同构。而这种同构性正是网络在基于全局信息攻击下如基于度的蓄意攻击下,表现出脆弱性的根本原因[8]。

4 结束语

本文系统地研究了拷贝模型的鲁棒性问题,对其在随机故障和多种蓄意攻击模式下的关键功能指标进行了模拟分析,发现拷贝模型的鲁棒性与其拷贝因子之间有着显著关系。具体而言,若拷贝因子越大,则拷贝机制导致的结构冗余规模越小,其相应网络对蓄意攻击越鲁棒,而对随机故障却越脆弱。本文的研究启发笔者进一步思考:拷贝或者复制机制对于生物系统鲁棒性的正面作用在网络中仅局限于基于有限信息的随机攻击模式;基于拷贝机制构造的结构冗余网络,还不足以有效抵抗基于网络全局结构信息的蓄意攻击。而人们在直觉中所理解的拷贝或复制对于网络鲁棒性所起到的正面作用,很可能体现在本文所验证的拷贝模型对于随机攻击的鲁棒性上。为此,笔者将进一步探索健壮网络的生成机制,使其在面临蓄意攻击和随机故障的风险时均能具有较好的抗攻击性能。

参考文献:

[1] WATTS D J, STROGATZ S H. Collective dynamics of small-world networks[J]. Nature, 1998, 393(6684): 440-442.

[2] BARABASI A L, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439): 509-512.

[3] RAVASZ E, BARABASI A L. Hierarchical organization in complex networks[J]. Physical Review E, 2003, 67(2): 026112.

[4] NEWMAN M E J. Mixing patterns in networks[J]. Physical Review E, 2003, 67(2): 026-126.

[5] BARAHONA M, PECORA L M. Synchronization in small-world systems[J]. Phys Rev Lett, 2002, 89(5): 054101.

[6] YANG Hui-jie, YIN Chuan-yang, ZHU Gui-mei, *et al*. Self-affine fractals embedded in spectra of complex networks[J]. Physical Review E, 2008, 77(2): 045101.

[7] XIAO Yang-hua, XIONG Mo-miao, WANG Wei, *et al*. Emergence of symmetry in complex networks[J]. Physical Review E, 2008, 78(6): 046102.

[8] ALBERT R, JEONG H, BARABASI A L. Error and attack tolerance of complex networks[J]. Nature, 2000, 406(6794): 378-382.

[9] ERDOS P, RENYI P. On the evolution of random graphs[J]. Publ Math Inst Hung Acad Sci, 1960, 5: 17-61.

[10] HOLME P, KIM B J. Attack vulnerability of complex networks[J]. Physical Review E, 2002, 65(5): 056109.

[11] COHEN R, EREZ K, BEN-AVRAHAM D, *et al*. Resilience of the Internet to random breakdowns[J]. Phys Rev Lett, 2000, 85(21): 4626-4628.

[12] PAUL G, TANIZAWA T, HAVLIN S, *et al*. Optimization of robustness of complex networks[J]. Eur Phys J B, 2004, 38(2): 187-191.

[13] TANIZAWA T, PAUL G, COHEN R, *et al*. Optimization of network robustness to waves of targeted and random attacks[J]. Physical Review E, 2005, 71: 047101.

[14] PAUL G, SREENIVASANA S, HAVLIN S, *et al*. Optimization of network robustness to random breakdowns[J]. Physica A, 2006, 370(2): 854-862.

[15] TANIZAWA T, PAUL G, COHEN R, *et al*. Optimization of the robustness of multimodal networks[J]. Physical Review E, 2006, 74(1): 016125.

[16] 张宁, 张丹荣, 杨建民. 有限资源条件下网络病毒的阻断策略[J]. 上海理工大学学报, 2007, 29(3): 250-254.

[17] 张丹荣, 张宁. 复杂网络下引入时间参数的病毒传播[J]. 微计算机信息, 2007, 23(36): 204-205.

[18] 朱刚, 张宁, 马良. 复杂网络上计算机病毒传播和控制策略研究[J]. 计算机应用研究, 2006, 23(9): 54-56.

[19] 谭跃进, 吴俊, 邓宏钟, 等. 复杂网络抗毁性研究综述[J]. 系统工程, 2008, 24(10): 1-5.

[20] 邓宏钟, 吴俊, 李勇, 等. 复杂网络拓扑结构对系统抗毁性影响研究[J]. 系统工程与电子技术, 2008, 30(12): 26-28.

[21] KUMAR R, RAGHAVANY P, RAJAGOPALAN S, *et al*. Stochastic models for the Web graph[C]//Proc of the 41st Annual Symposium on Foundations of Computer Science. Washington DC: IEEE Computer Society, 2000: 57-65.

[22] CHUNG F, LU L, DEWEY T G, *et al*. Duplication models for biological networks[J]. J Com Bio, 2003, 10(5): 677-687.

[23] XIAO Yang-hua, MACARTHUR B D, WANG Hui, *et al*. Network quotients: structural skeletons of complex systems[J]. Physical Review E, 2008, 78(4): 046102.