

复杂网络抗毁性优化研究综述

张煜¹ 张琨²

(1.信息系统工程重点实验室, 江苏南京210000; 2.南京理工大学计算机科学与技术学院, 江苏南京 210094)

摘要: 随着对复杂网络认识的不断深入, 复杂网络抗毁性的优化策略成为了新的研究热点。文章从拓扑结构、网络容量以及路由策略等三个层次详述了目前国内外复杂网络抗毁性优化策略的研究进展。最后对复杂网络抗毁性研究存在的问题进行了总结, 对未来的发展趋势作了展望。

关键词: 复杂网络; 无尺度网络; 抗毁性优化

中图分类号: TP393 **文献标识码:** A **文章编号:** 1003-9767 (2011) 06-0126-04

1. 引言

复杂网络广泛存在于自然界、生物界、工程界和人类社会界, 如交通网络、神经网络、蛋白质网络、Internet、万维网、电力网、朋友关系网以及金融网络等等。深入研究复杂网络, 可以揭示隐藏在自然界、生物界、工程界和人类社会大量复杂系统中的共同规律。这些一般性规律的揭示对于把握复杂网络的宏观特征, 调节复杂网络上的动力学行为都将具有重要意义。

随着国际学术界对复杂网络研究的日益关注, 国内已有学者开始这方面的研究并取得了一些阶段性成果, 但是到目前为止还较少有系统介绍国内外关于复杂网络抗毁性优化策略的综述文献。本文从复杂网络模型的演化入手, 在简要介绍复杂网络统计特征的基础上, 对国内外研究现状进行综述, 以期对国内关于复杂网络的研究起到进一步的推动作用。

2. 复杂网络概述

2.1 基本参数

目前的复杂网络研究还处于起步阶段, 大量的实证研究方法被用到复杂网络中来。不同领域的研究者通过统计的方法对复杂网络的研究表明, 要发现不同系统抽象出来的复杂网络的共同特征和衡量复杂网络拓扑模型的准确性, 就必须设定统一的复杂网络参数。因此, 在目前的研究中, 使用较为频繁的几个参数为网络的度分布 (degree distribution)、网络平均距离 (shortest path length)、网络聚集系数 (clustering coefficient) 以及介数 (betweenness) 等。

(1) 网络节点的度分布

在网络中, 节点的度是指与该节点相邻的节点的数目, 即连接该节点的边的数目。而网络的度 $\langle k \rangle$ 指网络中所有节点度的平均值。度分布 $P(k)$ 指网络中一个任意选择的节点, 它的度恰好为 k 的概率。

(2) 网络平均距离

从结点 i 到结点 j 的最短路径 l_{ij} 是指所有从 i 到 j 的连通的通路中, 所经过的其他结点最少的一条或几条路径。平均最短路径 L 是对所有 l_{ij} 的平均。

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} l_{ij}$$

(3) 网络的聚集系数

用 C_i 表示, 指的是在节点 i 周围存在 k_i 个节点, 这 k_i 个节点间存在的互相连接的数目为 E_i , 而这些边的数目同可能存在于这 k_i 个节点的边的数目之比称之为聚集系数:

$$C_i = \frac{2E_i}{k_i(k_i-1)}$$

它表示这个节点的连通程度, 是网络局部化特性的表现。如果对所有节点的聚集系数取平均, 则称之为网络的平均聚集系数。

(4) 介数

包括节点介数和边介数。节点介数指网络中所有最短路径中经过该节点的数量比例, 边介数则指网络中所有最短路径中经过该边的数

量比例。介数反映了相应的节点或边在整个网络中的作用和影响力。一个节点的介数越大, 流经它的数据分组越多, 意味着它更容易阻塞, 成为网络的瓶颈。

如果记图中任意两点 i, j 之间的最短路径条数为 σ_{ij} , 而这些最短路径中经过节点 w 的条数为 $\sigma_{ij}(w)$, 那么节点 i, j 之间经过 w 的最短路径条数占 i, j 间总的最短路径条数的比例为 $\sigma_{ij}(w)/\sigma_{ij}$, 在此基础上, 节点 w 的介数定义为:

$$C_g(w) = \sum_{i \in V} \sum_{j \in V, j \neq i} \frac{\sigma_{ij}(w)}{\sigma_{ij}}$$

2.2 几种演化模型

网络模型的发展历经了两大阶段, 第一阶段是以经典的图论知识为支撑, 对规则网络的研究。第二阶段以1959年Erdős和Rényi发表的《随机图论》为代表, 从此网络模型的研究进入了复杂系统时期。在第二阶段中又可分为三个时期, 它们分别以三个经典的网络模型为代表。其中1959-1998年以随机网络的ER模型为代表; 1998-2000年以小世界网络的WS模型为代表; 2000年到现在以无尺度网络的BA模型为代表。

(1) ER模型

在众多的无向无权网络模型中, 研究范围最广、时间最长的应该是Erdős和Rényi于1959年提出的ER(Erdős-Rényi)模型。ER模型的提出使网络模型研究进入到一个崭新的阶段。模型的构建规则如下:

设网络中共有 N 个节点, 每一对节点之间以概率 p 相连。在ER模型中, 网络最多可以容纳连接边的数目为: $N(N-1)/2$ 条; 若节点间以概率 p 相连, 它的实际连接的数目为: $pN(N-1)/2$; 则度的分布满足二项式分布

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

所以整个网络的平均度为 $\bar{k} = p(N-1)$ 。当 N 较大时, 上式的分布表达式将呈现一种Poisson分布的形式

$$P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!}$$

Poisson形式的度分布为ER模型的网络所特有, 已成为与其他网络的最大区别, 受此影响, 研究网络模型的许多学者大多都以网络的度分布作为区分不同网络的标准之一。ER模型网络的节点度服从Poisson分布的形式。

(2) Small-world网络模型

“小世界”网络特性是现实世界网络当中普遍存在的一种现象, 1998年Watts和Strogatz对这一现象进行了总结。Watts和Strogatz通过对ER模型的研究并结合自然界的小世界效应提出了著名的WS模型。WS模型的提出掀起了复杂网络研究的热潮。

Watts和Strogatz提出的WS模型来源于自然界的网络, 通过观察他们发现, 这些网络平均最短路径比较小, 服从 $\bar{l} \sim \ln(N)$, 可以同ER网络模型相比拟, 但该网络的聚集系数却比ER网络模型高出许多。

所以,为了描述具有这种动力学特性的网络,Watts和Strogatz构造出了同时具有相对较小的最短平均路径和较大聚集系数的WS网络模型,并称具有这样两种特性的网络为小世界网络。

(3) Scale-free网络模型

上世纪90年代以来,科学技术的进步使得研究大规模的实际网络结构成为可能,在计算机的帮助下模拟真实世界网络也成为可能。正是在这样一种条件下,人们发现真实世界当中的网络并不像Erdos和Renyi所描述的那样是基于一种完全随机的网络,它们在构造过程中基于某种规律且在网络动力学特性上具有ER模型网络所不具备的特性,在这些网络动力学特性中最典型的是小世界特性和节点度的幂律分布特性。

为了构造出具有无尺度特性的网络模型,各国的科学家开展了不懈的研究,其中尤以Barabasi和Albert最为杰出。他们通过对包括英特网、万维网和新陈代谢网络在内的许多大型网络拓扑结构的研究,提出了具有无尺度网络特性的BA网络模型,通过各国科学家的共同努力该模型已成为无尺度网络模型研究的基础。

1) 增长特性 整个网络开始于一个具有 m_0 个节点的初始网络,每隔一段时间间隔后向初始网络中增加一个新节点,同时生成 $m(m \leq m_0)$ 条与系统中已存在节点相连的边;

2) 择优连接 当新节点加入到网络的时候,以概率 $\Pi(k_i)$ 与网络中其它节点 i 产生连接,而该概率取决于节点 i 的度数,也就是说

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

网络在 t 个时间步这样的操作之后,将存在 $N = m_0 + t$ 个节点,和 mt 条新增加的边。根据增长性和倾向性选择,网络将最终演化成一个标尺不变的状态:网络的度数分布不随着时间而改变。

2.3 故障与攻击的承受性

在随机网络中若除去大部分节点,网络立刻瘫痪,但无尺度网络却对意外故障具有很强的鲁棒性,随机除去那些具有少量连结的节点,对网络影响不大,系统依然运行,故无尺度网络具有很强的鲁棒性。

虽然无尺度网络具有鲁棒性,但它同时也存在脆弱性,由于集散节点拥有大量的连接,其数量高达整个系统的80%,它支撑着整个系统的运行,如果除去集散节点或对集散节点恶意攻击,后果将不堪设想。例如,当集散节点遭受病毒感染和破坏时,病毒就会传播给众多的其他节点,由此会迅速感染整个网络系统。另外,电脑黑客蓄意攻击一些集散节点,也足以使整个网络系统瘫痪。无尺度网络对于恶意攻击的这种脆弱性犹如“阿喀琉斯之踵”,如果不对其抗毁性加以优化,后果将不堪设想。

因此,需要针对无尺度网络的这种双重特性提出有效的抗毁性优化策略,从而既保持其对意外故障的鲁棒性,又提高它抵抗选择性攻击的能力。

3. 复杂网络抗毁性优化策略

在复杂网络抗毁性研究的实际应用中,往往不是重新设计、构建一个网络,而是在已有网络的基础上,提出具体的抗毁性优化策略。研究复杂网络抗毁性的最终目的,就是如何用最少的成本,建设一个抗毁性好的复杂网络。国防科技大学的谭跃进教授等,把复杂网络抗毁性优化设计分为了三个层次,本文也从这三个层次入手,介绍近年来国内外在复杂网络抗毁性优化策略方面的研究进展。

3.1 网络拓扑结构的优化设计

作为提高网络抗毁性的第一个层次,拓扑结构的优化设计使得网络在发生故障的情况下重要业务不受影响,其主要手段就是在网络中增加足够多的备份链路和备份设备。通过增加网络中节点或边的数量,能缩短网络节点之间的最短路距离,提高网络连通性,从而当面

临打击时,网络可以通过其他节点和边保持工作能力。当然提高网络抗毁性也不是一味地增加备份链路、备份设备,因为如果在网络的任意两个节点之间都连接一条链路,这样的网络虽然具有极高的抗毁性,但也需要花费极高的网络建设成本。

(1) 解析法优化网络拓扑结构

Shargel等研究了参数可调BA模型上的抗毁性优化问题。他们引入两个可调参数 $p \in [0,1]$, $g \in [0,1]$,分别对应BA模型中的择优连接和网络增长的概率。通过仿真的方法优化参数组来优化网络的抗毁性,发现当 $(p,g)=(1,0)$ 时网络是最优的。

Paul等认为对于给定度分布的网络可能对某种类型的攻击具有强抵抗力。以提高网络对随机失效和选择性攻击的综合抗毁性为目标,得到了如下三种网络的优化参数:1) 无标度网络;2) 双幂率网络;3) 双峰分布网络。在这些网络分布中,当网络符合仅一个节点拥有极大的度 $k_2 \sim N^{2/3}$,而所有其他节点的度数为 $k_1 \approx \langle k \rangle$ 这样的双峰分布时具有相对最强的综合抗毁性。文献虽然综合考虑了随机失效与选择性攻击,但随机失效与选择性攻击并不是同时发生在网络上。

Tanizawa等把随机失效与选择性攻击描述成一系列的“攻击波”,并研究在这样的攻击波下网络抗毁性的优化。他们在一轮攻击波中选择性移除 pt 节点,随机移除 pr 节点,发现当比例为 r 的节点度数为 $k_2 = (\langle k \rangle - 1 + r) / r$,其余节点度数为 $k_1 = 1$ 时网络具有相对最优的抗毁性。此时网络的渗透阈值 f_c 最大,且 f_c 仅与 p_r/p_t 有关。

Wang等把网络面对随机故障的抵抗力优化转化为度分布熵的优化,利用熵优化模型,研究了无标度网络面对随机故障的抗毁性。他们发现,在无标度网络的最小度给定的条件下,网络对随机攻击的抵抗力随着网络规模的增加而增强,随着网络的标度指数的增加而减弱,随着网络平均度的增加而增强。Wang等进一步探讨了对随机攻击有抵抗力的网络设计的一般性原则,得出具有少量中心节点、高聚集系数的网络与随机网络相比具有更好的抵抗力。

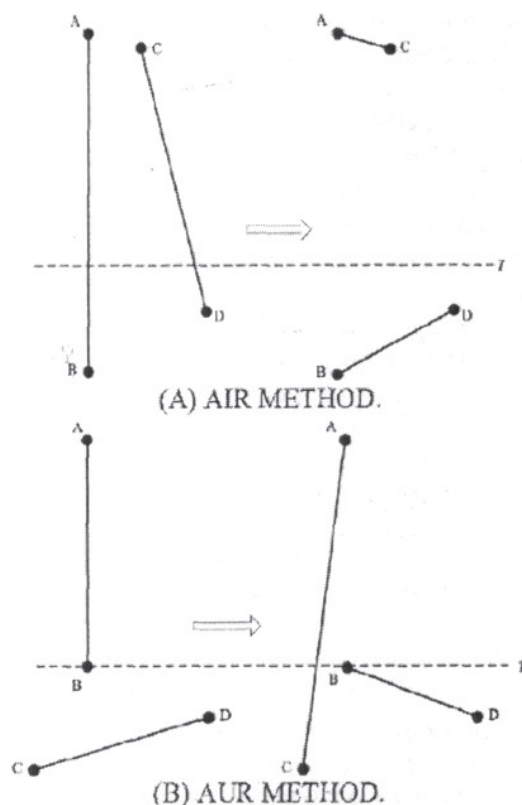


图1 两种重建策略AIR与AUR

(2) 增加/重建连接策略提高抗毁性

S.Xiao等在中证明了失去一定比例的中等大小或者中心节点,能够降低蓄意攻击的效率。此外,他们在文献中还提出了两种重建连接策略AIR与AUR,根据事先设置的阈值T,对高度数以及低度数节点之间的连接进行移除与重连。

如图1(a),假设节点A与C度数高于阈值T,而节点B与D度数低于T,AIR策略把A-B、C-D之间的连接替换为A-C、B-D。这种方法会增加网络的匹配混合度 assortative mixing,并且有可能导致网络中的一些节点断开连接。第二种重连方法AUR则不会引起匹配混合度的明显变化,它的主要思想是改变相对“低度数”节点之间的连接模式,如图1(b)。设置一个较小的阈值T,假设节点A度数的大于T,节点B的度数等于T,而节点C与D的度数均小于T。重连操作同样也是把A-B、C-D之间的连接替换为A-C、B-D。仿真实验证明,这两种方法都能在不增加匹配混合度以及平均路径长度的同时,增加网络的抗毁性。

Beygelzimer等通过添加边的方式来优化网络结构。他们提出了六种添加边的方式,并且通过仿真实验,得出其中四种方式的优劣顺序。但是,当攻击的节点数目超过一定的程度,不管重连的边数有多少,网络抵抗攻击的能力也不会明显增加。

(3) 遭受攻击后修复网络连通性

以上研究都是基于攻击同时发生,网络无法进行响应的基础上。例如随机故障可以看成是同时删除随机选择的一系列节点和边,网络无法进行任何防护和补救。而往往真实的攻击是一种攻击逐步进行的“渐进式”攻击,网络可以对攻击作出响应,即在遭受攻击时采取积极的补救措施以保持网络的完整性,这也是更为有效的抗毁性优化策略。因此,国内外的研究者提出了一些在网络遭受攻击后修复网络连通性的方法。

Rezaei等提出了补偿动力学、分布式治愈算法和高效创建极大连通分量三种方法,能够把选择性攻击所造成的危害降到最低。在由于高连通度节点被攻击而分裂成碎片的PL网络中,仅仅引入一小部分的随机边,通过“反渗透”进程,就可以重建连通性,进而能够重建原网络的其他拓朴性质。

概率为 p 的“反渗透进程”可以解释为引入随机边的过程,即以比例 p 在超图的所有连接中做随机配对。对任意的 k ,网络中将有 $NP(k)$ 个度数为 k 的节点,标记为 $I_1^k, I_2^k, \dots, I_{NP(k)}^k$ 。为这些节点中的每一个创建 k 个备用副本节点 $I_{i1}^k, I_{i2}^k, \dots, I_{ik}^k$,如图2(a)。然后在这张超图上就能够开始随机配对过程,如图2(b)。配对完成之后,把节点 i 的所有副本合并为一个节点,因此所有副本的链接现在都连接到了合并的节点上,如图2(c)。

相对国外而言,国内的研究起步较晚,但仍有不少学者进行了有益的探索。如华中师范大学的蔡勋教授等进行了遭受复杂网络的修复策略与关联特征研究,它以一定的修复几率将遭受节点和其它节点重连,来提高网络的抗毁性,并以美国航空网络的实际数据,分析并验证了所提方法的有效性。刘玉华教授等通过避免集散节点的形成来确保网络的鲁棒性。他们提出了两种避免集散节点形成的策略,一种是层次结构,一种是分布式结构。这两种结构中的节点之间具有良好的通信能力,一旦某个节点出了故障,则邻近的节点可以迅速的代替它或是把它隔离。这两种结构都具有灵活性与容错性,仿真结果显示,这两种结构比BA模型中的集散节点具有更好的鲁棒性,网络系统可以更正常的运行。

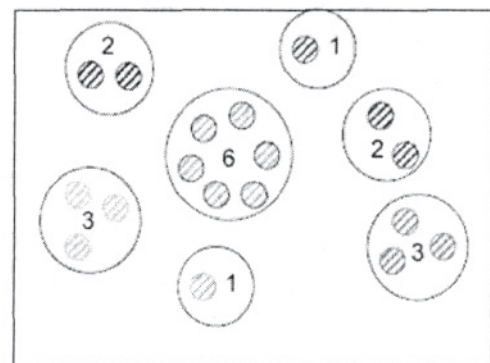
电子科技大学的隆克平教授等,考虑了在“协同攻击”情况下复杂网络的抗毁性问题。他们认为协同攻击是最具有破坏性的攻击之一,它能够协同所有已崩溃节点的信息,进而确定接下来的攻击目标。针对这种攻击,他们引入了R.C.H策略,移除一定比例的崩溃节

点,仿真结果证明此策略能从一定程度上增强网络的抗毁性。

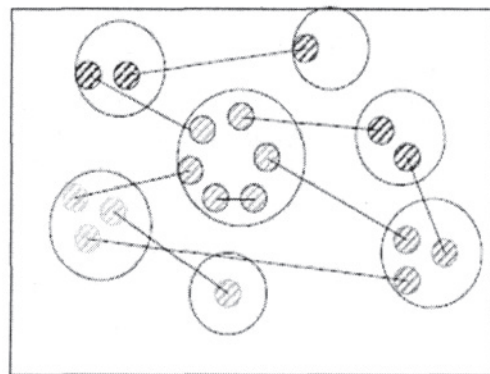
华东师范大学的刘啸林针对证券网络提出了使用生成树优化算法来求解某些抗毁性网络设计的难题。生成树优化算法解决了最小连通度为2、任意两节点间最大跳数不超过K的抗毁性网络设计问题,对大型网络的规划、设计或对已有网络进行扩容、优化具有一定实际应用价值。

3.2 网络容量的优化设计

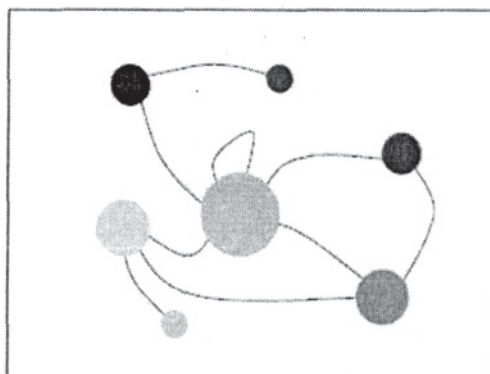
一旦网络的拓朴结构具备了一定的抗毁性的基础,在网络中出现故障后,网络仍然能够保持连通。接下来要考虑的就是网络容量优化问题。网络容量是影响网络性能的重要因素,在网络流的传递过程中,极有可能因为某些关键节点或边的容量限制而导致网络阻塞,进而引发导致全网崩溃的“级联失效”。例如一个连通度为2的网络,在任意一对节点之间一定存在不止一条的传输路径,如果一条传输路径出现故障,这对节点之间的传输流量就会转移到另一条传输路径上。所以在设计网络中每条链路的带宽时,就要考虑到必须有足够的



(a) 节点度数为 i , 创建 i 个副本



(b) 执行随机配对



(c) 副本合并成为一个节点

图2反渗透进程

空余带宽来容纳可能增加的额外流量。

一个具有一定抗毁性的网络,在任意一对节点之间一定存在不止一条的传输路径,如果一条传输路径出现故障,这对节点之间的传输流量就会转移到另一条传输路径上。因此在设计网络中每条链路的带宽时,就要考虑到必须有足够的带宽来容纳可能增加的额外流量。这一类问题的研究在WDM网络、ATM网络等光通信网络的应用中有着较大的实用价值。

中科院的范晶等引入了中介中心性的概念,基于此概念对网络拓扑进行优化和拥塞预测。通过理论分析和仿真实验,考察了一系列网络参数与网络容量的关系,最终提出依据网络中节点的介数以及介数的标准差增加一些捷径路径的方法,从而有效平衡中枢节点的负载,缓解拥塞状况,提高网络容量。

3.3 路由策略的优化设计

好的路由策略是网络持续发挥作用的基础。很多网络优化问题被证明是NP-hard的,因此,神经网络、遗传算法、禁忌搜索等启发式算法被广泛应用于研究这类问题。根据传输路径切换方法的不同,这类问题又分为静态路由策略和动态路由策略两种。

(1) 采用静态路由技术

当网络中出现故障时,在故障部分的两端节点上,把所有通过故障部分传输的数据切换到另一条传输路径上,从而绕过发生故障的部分。这另一条传输路径以及路由配置都是预先设定好的,所以称为静态路由技术。由于这类问题的计算复杂度较高,近年来有较多的研究者尝试用各种启发式搜索算法来求解这个问题。于波等尝试将遗传算法应用在网络流量及带宽分配中,以最小化全网平均分组时延和链路成本为目标函数,在得到流量分配方案的同时,求得最优链路带宽。申健等采用禁忌搜索算法来求解通信网络中的链路容量与流量分配(CFA)的问题,经过大量的仿真试验结果验证了禁忌搜索算法对求解CFA问题的有效性以及解的高质量性。

(2) 采用动态路由技术

与静态路由技术不同的是,当网络中出现故障时,是由传输业务的源端根据网络的拓扑结构以及链路上的备份带宽,实时地在网络中选择一条到达终端的传输路径。采用这种技术的优点是不需要预留带宽,可以有效利用网络的空余传输能力,缺点是网络上必须有足够的

备份带宽,而且对路由选择时间的要求比较高。

Harjani A等指出,在这种方式下,不同的通信线路保护策略各有优缺点,采用什么样的线路保护策略是需要在网络整体优化程度与故障恢复时间之间进行权衡的。为了寻找更快速、有效的路由搜索算法,郭磊等研究了WDM网状网中双链路失效的动态共享链路保护设计问题,提出了一种新的启发式算法TBP(Two Backup Paths)和一种改进的启发式算法OBP(One Backup Path)。两种保护算法都能根据当前网络状态动态地调整链路代价,所选路由都是最小代价的路径。

4. 结束语

复杂网络理论的研究目前已经发展成物理学、控制科学、生物学、管理学等几乎所有学科的前沿热点问题,复杂网络抗毁性研究也日益引起更多的注意。然而,目前的研究大多处于“认识”阶段,即通过统计物理学、动力学等途径对复杂网络抗毁性进行分析与评估,在优化策略方面的研究成果却相对较少。

因此,未来对复杂网络的研究将必然从“认识”阶段逐步向“优化”阶段转变,以交通网络、生物网络、社会关系网络等实际网络为背景,对抗毁性进行建模、分析、优化、控制,提高它们对蓄意攻击的抵抗力。在优化策略方面,将不仅仅局限于对拓扑结构的优化,网络的最优防御策略、最优故障修复策略等将成为重要研究方向。另外,在现有文献中进行节点失效研究时,很少考虑节点本身具有一定的主动防御能力和自我修复能力。在网络模型中综合考虑上述因素会得到怎样的结论,是值得探讨的问题。

参考文献:

- [1] 王林,戴冠中. 复杂网络的Scale-free性、Scale-free现象及其控制[M]. 北京: 科学出版社, 2009.
- [2] Adamic L A, Huberman B A. Power-law distribution of the world wide web[J]. Science, 2000, 287:2115.
- [3] Maslov S, Sneppen K. Specificity and stability in topology of protein networks[J]. Science, 2002, 296: 910-913.

基金项目: 南京理工大学自主科研项目, 项目批准号: 2010ZYTS035; 南京理工大学紫金之星项目, 项目批准号: 2010ZJZX012.

(上接第125页)

员同志,选择某个部门名称可以查看属于该部门的所有教职工,方便进行统计查询。

4. 结束语

高职院校人事信息管理系统的网络化具有许多优点,例如:系统设计先进,功能齐全,界面友好,使用方便,适用性强,安全性高等等。目前,系统正在积极地开发,待系统投入运行后,将使高职院校人事部的教职工从以往复杂、简单的劳动中解脱出来,以更多的精力投入到人事决策以及高职教学管理工作当中。

参考文献:

- [1] 徐慧军. 简述企业人事信息管理系统的设计[J]. 知识经济, 2010,

(07).

- [2] 李娜,王春梅,武春杰. 企业人事管理系统的设计与实现——查询模块的实现[J]. 沿海企业与科技, 2009,(03).
- [3] 丁磊,孙佳帝. 基于C#编程技术的企业人事管理系统的设计与研究[J]. 才智, 2010,(08).
- [4] 周新邵. 基于ASP技术的网上高校毕业生档案管理系统的设计与实现[J]. 湖南科技学院学报, 2007,(04).
- [5] 谭国臣,王晓涛,娜嘉,刘树华,于建永. 高校人力资源管理信息系统的开发[J]. 鲁东大学学报(自然科学版), 2007,(03).
- [6] 李娜,王春梅,武春杰. 企业人事管理系统的设计与实现——查询模块的实现[J]. 沿海企业与科技, 2009,(03).