

基于复杂网络的网络业务流建模与行为分析

摘 要

作为社会信息化发展的基石,通信网络的核心价值在于网络中承载的业务,探索业务行为特征能够聚焦于业务本身特点,发现业务的内在传输规律。现代通信网络日益复杂、智能化,与此同时,海量业务共存于网络使得业务耦合性、突发性不断增强,业务服务质量难以得到有效保障,动态复杂网络环境下基于业务服务质量的行为特性研究成为亟待解决的问题。目前关于业务行为的研究大多集中于业务生成、到达及分布规律,一方面很难直接与业务服务质量相关联,另一方面研究成果大多基于稳定网络环境,具有一定局限性。针对已有研究存在的不足,本文重点研究复杂通信网络中业务服务质量间的耦合关联机理,挖掘网络突发和故障环境下业务的传播动力学行为,并提出有效的传播控制方案,保障业务安全可靠传输。具体而言,本文的创新点主要有以下四个方面:

(1) 提出了一种基于弹性力学的业务流耦合关联网络模型,挖掘业务服务质量间的耦合机理。

针对网络拥塞环境下业务服务质量表现出的普遍关联现象,本文引入弹性力学理论构建了无向加权的业务流耦合关联网络,用于解决大规模复杂网络中的业务关联性问题。该网络中节点表示通信业务,连边表示业务服务质量之间的弹性关系,通过对其结构特性进行仿真,发现当业务随机分布并达到一定数目时表现出“小世界效应”,即业务服务质量间存在紧密、集团化的耦合关系;网络中业务节点的度分布并不均衡,存在少量度数很大的业务流,对全网业务服务质量具有广泛的影响作用。

(2) 提出了一种突变业务流影响传播模型,分析其传播机理与传播规律。

针对突变业务流影响传播机制建模问题,本文提出了一种基于

耦合概率与邻居关系的突变业务流影响传播模型，分析不同网络结构、不同网络状态下突变流的传播范围、传播强度规律。仿真结果表明，业务突变强度、网络拥塞程度、网络结构类型均会对传播效果造成影响，并遵循统一的规律：突变流的传播强度随着业务邻居阶数增长而大幅下降，下降幅度与各阶邻居业务的数目相关，且对奇数阶邻居与偶数阶邻居的带宽影响相反。

(3) 提出了一种基于业务社团结构的免疫控制算法，保障业务安全可靠传输。

针对突变业务流引发的全网业务带宽波动问题，本文提出了一种基于业务社团结构的免疫控制算法。传统方法大多通过改变流量生成或传输的方式来削弱业务流突变强度，增加了系统开销，降低了业务的传输效率。本文提出的算法通过以社团为单位，免疫社团边界业务流以切断突变流传播路径，在不改变大部分业务正常传输的情况下，极大地缩小了突变流影响范围。仿真结果表明，该算法最优可达到利用 21% 的免疫业务保护 97.2% 的业务免受带宽波动影响。

(4) 提出了一种基于双层网络的级联故障模型，分析业务流动态行为特征。

针对网络负载不均衡导致的通信链路失效问题，本文建立了一种基于业务层/网络层双层网络的级联故障模型，研究故障中业务传输行为与竞争行为的动态变化过程。仿真结果表明：1) 级联故障中业务对网络的依赖强度与业务平均最短路径长度成正比，间接反映了业务传输效率的高低；2) 网络最大连通片可反映业务竞争范围的变化，在蓄意攻击下，业务竞争范围受级联故障的影响比随机攻击更为明显；3) 业务竞争强度的变化与业务度分布和拥塞程度相关，网络中业务分布的不均匀性会导致业务竞争强度更易受到级联故障的影响。

关键词： 业务行为 复杂网络 耦合关联 传播控制 级联故障

NETWORK TRAFFIC FLOW MODELING AND BEHAVIOR ANALYSIS IN COMPLEX NETWORKS

ABSTRACT

The core value of communication network lies in the services it carried due to the significance of social informatization. Flow behavior analysis focuses on flow characteristics and help us understand network regulations. Nowadays, networks are more and more complex and intelligent, at the same time, the coupling and bursty characteristics of flows are increasingly enhanced, flow behavior analysis becomes a difficult problem. By now, most studies focus on flow generations, arrivals, and distributions. On one hand, it can hardly build relations to flow transmission quality, on the other hand, these results are mostly based on a stable network environment which have limitations to some extent. Considering these deficiencies, this paper focuses on flow coupling mechanisms in complex communication networks, explores flow propagation dynamics in network bursty and failure environments, and proposes effective propagation control schemes to ensure flow reliable transmissions. The innovations and contributions of this paper are as follows:

(1) A flow interaction network based on elastic mechanics is established to explore the coupling mechanisms of flow transmission qualities.

To solve the QoS coupling problem, an undirected weighted flow interaction network is constructed by utilizing elastic mechanics, which is applicable to flow correlation analysis in large-scale complex networks. In this network, nodes denote flows, edges with weight denote elastic relations between flows. By simulating the structural characteristics of the network, ‘small-world effect’ was found when flows randomly distributed

and reaches a certain quantity, which indicates the close and clustered coupling relationship of flows. The flow degree distribution is uneven, there exists a small portion of flows with very high degrees, which have great influence on flows' transmission quality.

(2) An influence propagation model of Internet bursty flows is established to analyze its propagation mechanisms and regularities.

An influence propagation model based on coupling probability and neighbor relationship was proposed to study the propagation characteristics under different network types and situations. Simulation results show that flow bandwidth bursty strength, network congestion level, and network types will affect the propagation process, but all follows a unified rule: the influence propagation of the bursty flow is a process that greatly decreases as the neighbor order increases due to the dispersion effect in each order, and the influence on odd-order neighbors and even-order neighbors alternates positively and negatively.

(3) An immune control strategy based on flow community structure is proposed to ensure reliable flow transmissions.

To control the transmission bandwidth fluctuations caused by network bursty flows, an immunization strategy based on flow community structure was proposed. Most of the traditional methods control the flow bursty strength by increasing the system overhead which reduce the flow transmission efficiency. Without changing the flow transmission modes, our algorithm greatly reduces the influence scope of the bursty flow by cutting off the propagation path based on the flow community structure. Simulation result shows that 21% immunization flows can protect 97.2% flows from bandwidth fluctuations in an optimal situation.

(4) Network cascade failure model based on coupled networks is established to analyze the dynamic behavior of flows.

To understand the cascading failures caused by unbalanced traffic load, a cascading failure model based on flow layer/network layer

coupling network was established to study the flow transmission and competition behavior. Simulation results showed that: 1) the dependence of the flow on the network is proportional to the flow average shortest path length, which reflects the flow transmission efficiency; 2) the network giant component can reflect the flow competition area, which under targeted attacks is much more obviously affected compared with random attacks; 3) the flow competition intensity is closely related to flow degree distribution and flow congestion degree. The uneven flow distribution in the network will obviously affect the flow competition intensity in cascading failures.

KEY WORDS: flow behavior complex network coupling correlation
propagation control strategy cascade failure

目录

第一章 绪论.....	1
1.1 研究背景与意义	1
1.1.1 网络规模庞大复杂化	1
1.1.2 多元业务爆发式增长	3
1.1.3 选题意义	4
1.2 研究现状	5
1.2.1 网络业务流行为定义	5
1.2.2 业务流行为研究现状	6
1.2.3 已有研究存在的不足	9
1.3 研究内容及关键问题	10
1.4 论文组织安排	12
1.5 本章小结	12
参考文献	12
第二章 复杂网络相关理论.....	19
2.1 网络及其表示方法	19
2.2 复杂网络特征度量	20
2.3 典型复杂网络结构	23
2.4 通信领域复杂网络研究进展	27
2.5 本章小结	29
参考文献	29
第三章 基于复杂网络的业务流服务质量耦合行为建模与机理研究.....	33
3.1 业务流弹性物理模型	34
3.1.1 弹性理论引入	34
3.1.2 模型建立	36
3.1.3 模型证明	40
3.2 业务流耦合关联网路	44

3.3 实验仿真与结果分析	45
3.3.1 网络环境与参数设置	45
3.3.2 业务流小世界特性	46
3.3.3 业务流拥塞行为评估	49
3.4 本章小结	51
参考文献	52
第四章 基于业务流突变的服务质量波动传播特性与控制算法研究	55
4.1 基于耦合关联的突变业务流影响传播模型	55
4.1.1 传统传播模型	56
4.1.2 突变业务流影响传播模型	58
4.1.3 模型精确度	63
4.1.4 实验仿真与结果分析	65
4.2 基于社团结构的突变业务流免疫控制算法	72
4.2.1 算法设计	73
4.2.2 实验仿真与结果分析	76
4.3 本章小结	83
参考文献	84
第五章 基于流过载失效的网络故障传播模型与业务流动态行为研究	89
5.1 渗流理论基本原理	89
5.2 基于流过载的双层网络级联故障模型	91
5.2.1 业务层/网络层双层网络模型	91
5.2.2 流过载级联故障模型	93
5.2.3 故障中业务动态行为模型	96
5.3 实验仿真与结果分析	98
5.3.1 网络环境与参数设置	98
5.3.2 网络级联失效	100
5.3.3 业务依赖强度（传输效率）	102
5.3.4 业务竞争范围	102
5.3.5 业务竞争强度	104
5.4 本章小结	106

参考文献.....	106
第六章 总结与展望.....	111
6.1 主要研究成果.....	111
6.2 主要创新点.....	112
6.3 未来研究方向.....	114
参考文献.....	114
附录 1 缩略语.....	117
致谢.....	119
攻读学位期间发表的学术论文目录.....	121

第一章 绪论

1.1 研究背景与意义

随着现代通信网络的大规模部署，网络结构愈加复杂化、异构化；同时，网络用户数目与日俱增，业务类型复杂多元，使得有限的分布式网络资源难以充分调度满足爆发式增长的网络业务需求。大规模的复杂网络环境激化了网络资源与业务需求之间的矛盾，不仅造成了网络数据测量的困难，而且致使网络业务流行为更加难以捕捉和刻画，业务流行为规律挖掘面临新的挑战。

- (1) 现代通信网络规模庞大、结构复杂，使得业务流在传输方式、传输效率、传输质量等方面都体现出与以往不同的特点，成为“网络业务流行为研究”的第一大挑战；
- (2) 随着移动互联网的普及，网络业务爆发式增长，大量业务相互耦合，导致网络业务流行为模式不断发生变化，成为“网络业务流行为研究”的第二大挑战。

应对通信网络新挑战，探索网络业务行为新特征，不仅有助于发现网络的运行规律，而且可以优化网络资源分配，保障网络业务的服务质量，对提升网络整体性能具有非常重要的意义。

1.1.1 网络规模庞大复杂化

人类进行通信的历史已久，古时人们通过语言、壁画、书信等方式传递信息。随着科学技术的发展，19 世纪中叶，电报、电话相继发明，使信息传递可以用电信号作为载体，通信领域产生了根本性的变革。以电路交换技术（Circuit Switching）^[1]为核心的电话网是早期最重要的通信网络之一，电话网的发展满足了公众语音业务的需求，距今已有百年历史。电路交换通信过程中双方独占电路，虽然保障了业务的低时延，但电路接续时间长，线路利用率和通信效率极低。随着计算机的普及，数据业务需求逐渐显现。以分组交换技术（Packet Switching）^[2-3]为核心的计算机网络应运而生，并以惊人的速度蓬勃发展。人们通过计算机网络可以进行即时通信、文件传输、收发电子邮件、视频点播，通信方式的改变深刻影响着人类生活。近年来通信技术的成熟与网络部署规模的扩张，使得通信

网络向复杂、灵活、智能的方向发展，网络固有问题也日益凸显，如网络业务服务质量保障、网络安全等。因此，未来通信网络中业务流行为规律的挖掘面临非常严峻的挑战。

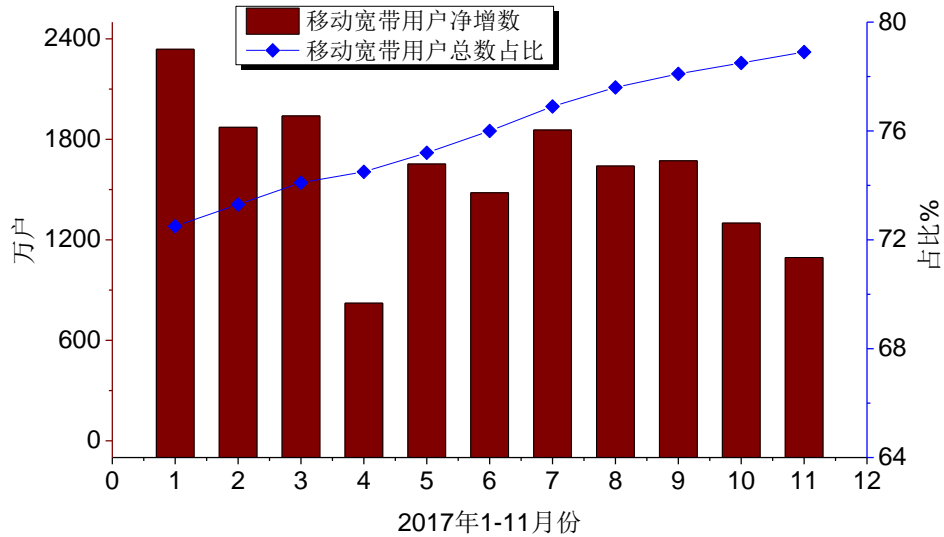


图 1-1 2017 年我国移动宽带用户增速和占比情况^[4]

首先，随着网络用户的增加，网络业务量急剧上涨，网络拓扑不断扩大，给网络业务性能参数的测量和分析造成巨大的困难。工信部公布的《2018 年 2 月份通信业经济运行情况》^[4]显示，截至 2018 年 2 月我国固定宽带接入用户达 3 亿户以上。大规模的宽带接入促使网络规模和承载能力相应增长。此外，随着 4G 网络服务的迅速普及^[5-8]，移动终端数目急剧增加，全民进入移动互联网时代^[9-10]。据统计^[4]，我国移动电话用户总数已超过 14 亿，其中 4G 用户接近 70%，且保持高速增长的趋势。图 1-1 显示了 2017 年我国移动宽带用户的增速和占比情况，截至 2017 年 11 月，移动宽带用户占比接近 80%，平均每月新增入网用户达到 1500 万户以上。

其次，通信网络规模的迅速扩大、演变，使得通信网络结构变得极为复杂，导致网络业务流在传输性能、传输效率等方面表现出不同的行为特征。传统通信网络的拓扑结构有很多种，例如星型结构、环型结构、总线结构、分布式结构、蜂窝结构等。随着网络规模的增长，通信网络结构逐渐呈现出新的规律。近年来，复杂网络^[11-14]理论蓬勃发展，Internet^[15-16]、WWW^[17-18] 等网络被发现具有无标度特征和小世界特征。无标度网络^[19]（Scale-Free network）的度分布符合幂律分

布，具有明显的异质性：网络中少数节点具有极高的度数，但大多数节点的度数却非常小。小世界网络^[20]（Small-World network）的特征路径长度小，一般稍高于随机网络；聚集系数高，可与规则网络相比较。通信网络的无标度特性和小世界特性直接关系到业务行为模式的变化，尤其是路由方式。

1.1.2 多元业务爆发式增长

（1）业务量级爆发式增长

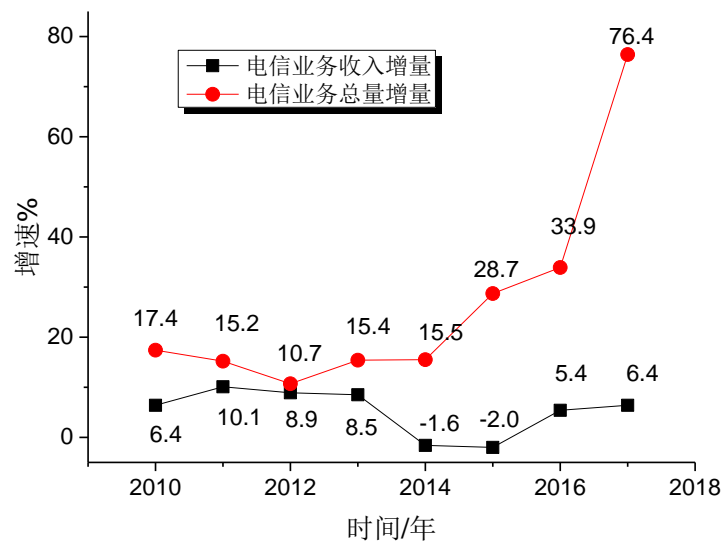
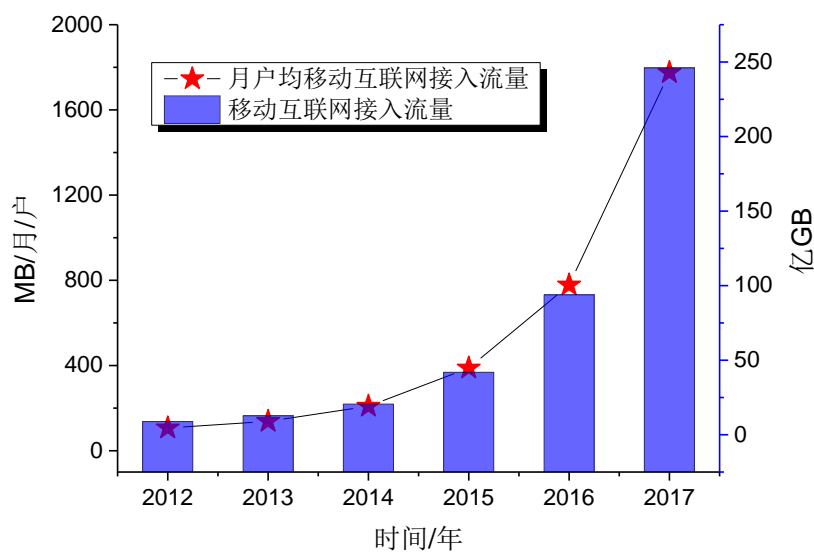
工信部 2018 年 2 月 2 日发布的“2017 年通信业统计公报”显示^[21]，2017 年，我国通信业在移动互联网、IPTV、物联网等方面取得了极大的成果。初步核算（图 1-2），2017 年电信业务总量达到 27000 亿元以上，相比于 2016 年超出 76.4 个百分点，增幅是 2010 年的 4 倍以上。电信业务收入约为 13000 亿元，比去年增长 6.4 个百分点，增速稳步提升。此外，我国移动互联网近年来呈现井喷式发展，4G 网络的普及更是使得线上支付、视频直播等数据流量爆发式增长。如图 1-3 所示，2017 年移动互联网接入流量达 246 亿 GB，约为 2016 年的 2 倍。全年月户均移动流量达到 1800MB，为 2016 年的 2 倍以上。

通信网络业务总量的大幅提升使得网络业务紧密耦合关联，网络任何运作方式的调整（网络结构、通信协议、资源分配），都会引发网络业务的大规模动态变化，并表现出与以往不同的行为规律。

（2）业务多元化

随着网络技术的进步，网络承载的业务从单一的语音业务逐渐向多元化的数据业务转变。尤其是近年来多媒体业务的兴起，网络业务呈现出多样化和综合化的趋势，且不同类型数据业务对服务质量有差异化要求。如何在复杂的网络环境下保障用户服务质量成为摆在人们面前的巨大挑战。

传统数据业务主要包括 web 访问、电子邮件、文件传输、远程登录、音视频点播、电话/视频会议等。而对于新兴多媒体业务，又增加了互联网业务、即时消息业务、移动商务业务、网络直播业务等。文件传输、电子邮件、消息发送等业务对于带宽需求和时延要求并不高，但传输可靠性十分重要；而视频类业务对带宽、时延的要求却很高，高带宽、低时延才能保障用户良好的上网体验。多元的业务需求使得网络对待不同的业务类型时需提供不同的保障方式，路由策略、排队机制、带宽分配原则的不同也将导致网络业务流行为模式间的差异。

图 1-2 2010-2017 年电信业务收入与总量增长图^[21]图 1-3 2012-2017 年移动互联网流量增长图^[21]

1.1.3 选题意义

随着通信网络的发展，一方面网络结构呈现出庞大复杂的特点；另一方面网络业务呈现出爆发式增长、多样化和综合化的趋势，网络业务间的耦合作用更为复杂。面对新的网络环境，挖掘业务流行为规律、维护网络高效运行成为摆在面前的巨大挑战。探索网络业务流呈现的行为特征，不但可以为网络高效运作提供

有效的数据资料和决策依据，而且能够聚焦于业务本身特点，针对性地解决服务质量提升问题。具体而言，网络业务流行为分析有以下优点：

- 服务质量保证

通过网络业务流行为分析，了解不同类型业务的传输需求与传输特征，采取差异化的服务方案，高质量保障业务可靠传输。

- 网络安全

预判可能出现异常的业务以及承载业务的链路和节点。管理员基于异常状况可以快速判定是否为网络攻击，并采取保护措施。甚至随着网络智能化的发展，可自动由系统判定故障类别、危险等级，并做出及时响应。

- 网络资源优化

分析不同时刻、不同地区的网络业务分布差异和网络资源利用情况，有效提高网络流量的负载均衡，促使网络提供时空差异化服务，提高网络资源利用率。

- 网络性能预测

预测网络流量及其变化趋势，如不同时段不同链路流量的分配情况，为未来网络的扩容、升级和整体规划提供决策依据。

1.2 研究现状

网络业务流行为研究逐渐广泛应用于通信网络性能分析、故障预测等领域，对于通信网络系统的构建、网络资源的优化利用、用户差异化需求的满足有重要指导作用。为了探索网络业务流行为产生和相互作用的机理，深层次挖掘内在原因，大量学者开展了相关研究工作，并取得了突破性的进展。

1.2.1 网络业务流行为定义

在通信系统中，网络行为分为用户行为和网络自身行为。用户行为由五项元素构成：时间、地点、人物、交互及交互的内容，具体而言包括用户习惯、用户喜好、用户分布等。网络自身行为分为网络运行、网络管理及网络维护行为，本文主要研究网络运行过程中业务流的行为特性。

网络业务流是指在网络中传输的一系列数据包的集合，主要包括 TCP 流、UDP 流、ICMP 流等。TCP/IP 协议族下的 TCP 流具有端到端的性质，且能保障

业务传输可靠性，是通信网络中非常重要的业务流类型。因此，本文主要考虑 TCP 业务流的行为特性。

所谓业务流行为特性，是指业务流在网络传输过程中体现出的一系列静态、动态、已知或未知的行为属性。本文将网络业务行为分为两类，一类是业务已知行为，另一类是业务未知行为。

“业务已知行为”是指业务在传输过程中可以提前预知的行为，如业务通信协议、路由算法、排队方式、带宽分配策略等。业务的已知行为可预知，是探究业务未知行为的基础。“业务未知行为”是指业务流在传输过程中表现出频繁变动、规律性或较难预测的活动，如业务对资源的竞争、合作行为，故障、突发环境下的动力学行为等。

本文主要基于网络业务的固有已知行为特性挖掘业务未知的行为规律，并依据规律对业务行为进行预测和有效控制，保障网络安全可靠运行。

1.2.2 业务流行为研究现状

大量学者在网络业务流行为研究方面取得了开创性的成果，主要分为以下四类：流到达规律研究、流长分布研究、流速特征研究和流行为应用研究。

（1）流到达规律研究

关于业务流到达时间间隔分布规律，目前学术界主要存在以下几种观点：业务流具有短相关性、具有长相关性和具有自相似性。

1) 具有短相关性 (Short-Range-Dependent)

针对业务的短相关性，学者们主要围绕泊松模型、马尔可夫模型和自回归模型^[22-25]展开研究。泊松模型 (Poisson Process) 假设业务流相互到达的时间间隔是相互独立的，满足泊松过程。Han 等人^[26]基于 SDN 网络提出了一种基于泊松散粒噪声过程的 TM (Traffic Matrix) 估计方法，简称为 FLAME。为评估 FLAME 算法预测效果，文献^[26]将估算结果与真实的数据中心网络流量进行跟踪比较。结果显示，FLAME 算法与实测流量相比相似度为 66.36%。

事实上，业务到达的时间间隔通常具有一定关联性，马尔可夫模型 (Markov Process) 利用时间序列的相关性，使业务到达的时间间隔产生相互关联，更符合实际通信网络业务情况。Ko 等人^[27]提出了一种利用时间序列和几何相关性的业务流量预测方法。该方法设计了三维热度图来描述链路状态，可有效地表征空间

和时间相邻的链路相关性。结果表明，所提出的方法能够以 85% 的精度预测流量状况，并且该准确度可以进一步提高。

自回归模型（Autoregressive Process）同样基于相邻业务的关联关系，将当前业务与下一业务之间建立了函数关系，依据这一特点，自回归模型成为最早运用在视频业务压缩的模型之一。李文璟^[28]等人提出了一种无线网络中基于 S-ARIMA 模型（Seasonal -Autoregressive Integrated Moving Average）的业务量预测算法，该算法在预测值和置信区间上都具有良好的结果。Wen 等人^[29]利用业务自相似性和重尾特性提出一种有效预测长程突发业务的自回归模型。

2) 具有长相关性（Long-Range-Dependent）

科学家通过深入研究发现了业务流的长相关特性^[30]，即业务流的到达在长时间范围内都具有相互关联关系。TES（Transform Expand Sample）模型是一种典型的长相关模型^[31]，可产生具有任意边缘分布和大范围自相关函数的自相关序列，但 TES 模型建模过程复杂，限制了其广泛应用。Bregni 等人^[32]详细比较了与流量长相关特性相关的各类 Hurst 参数的估计精度，如方差时间图、R / S 统计量、小波对数标度图等，对于有效选择 Hurst 参数进行流量预测具有重要帮助。

3) 具有自相似性（Self-Similarity）

Leland 等人^[33]对 Internet 流量的自相似性进行了详细的分析，证明了网络流量的自相似性是普遍存在的，后续学者的研究同样也显示自相似与分形模型更适合描述当前的网络业务^[34-35]。自相似模型生成业务流的过程主要有分形高斯噪声 FGN^[36-38]（Fractional Gaussian Noise）和分形布朗运动 FBM^[39]（Fractional Brownian Motion）两种方法。利用业务流量的自相似属性，文献^[40]提出一种有效预测网络突发流量的算法，并将常用的最小二乘法替换为能够更准确反映网络流量趋势的协变正交法，解决最优输出权重问题。通过分析视频流量轨迹，该算法可以更有效地捕获流量突发性的特征，相较于其他预测方案准确性更高。

除此之外，还有利用 M/G/∞ 排队模型^[41]和小波变换^[42]进行业务流到达规律研究的方法。

（2）流长分布研究

在流长分布方面，Web 流量是网络中主要的流量之一，研究主要集中在通过 Web 文件长度的分布推测 TCP 流的流长分布特征。Crovella 等人^[43]通过提取 Web 文件，发现 Web 文件长度服从重尾分布。在此基础上，Downey 研究了大量

来自不同服务器的 IP 流数据,发现 IP 流量的尾部特征服从 Lognormal 分布或者 Pareto 分布^[44],并推测 IP 流的自相似特性可能与其流长分布的尾部特征有密切关系。以上结论均通过真实网络数据分析获得,随着通信网络的飞速发展,模型的局限性逐渐体现:实验结果是否依然符合大规模复杂网络环境,以及是否依然满足爆发式增长的多元业务现状等。

(3) 流速特征研究

流速是指单位时间到达报文数或比特数的大小,研究通常包括业务流持续到达时间、单位时间平均流速和稳定性三个方面。流速特征的研究可以为路由策略、流量建模和网络安全等方面提供决策依据,目前代表性学者有 Sarvotham^[45]、Brownlee^[46]、Lan^[47]、Park^[48]等人。Sarvotham^[45]改进了传统的 on-off 模型,提出了 $\alpha-\beta$ 模型用来分别描述网络高速率大文件和低速率小文件,该模型能帮助更好地理解网络流量突发性和长相关性产生的原因。此外,文献^[45]针对互联网流量动态分析、拥塞控制、真实流量模拟所需的关键参数提供了新的见解。Brownlee^[46]等人发现,网络中约 45% 的业务流持续时间少于 2 秒,而 50% 至 60% 的大量业务流生命期可以达到数小时甚至数天。因此,网络设备存储配置方面必须能够处理大量短流,并依据业务短流和长流之间行为差异,改善路由器的缓存转发机制。

(4) 流行为应用研究

近年来许多学者在业务流行为应用方面取得了实质性的发展。研究主要包括基于流行为特征的业务流分类方法、基于流行为的网络异常检测等。通过提取业务流相关信息,例如用户行为习惯、通信协议、路由规则以及传输过程中参数呈现的特点,对业务流特征进行识别,实现基于机器学习的网络流分类与故障检测。

在业务流分类方面,Moore 等人提出了基于朴素贝叶斯 (Naive Bayes) 的流分类方法^[49],统计得到上百个流特征,并利用不同特征组合对应的业务类型不同,进行机器学习训练,最终实现业务的精确分类。Huang 等人提出了基于 K-最近邻分类算法 (K-nearest Neighbor)^[50],将业务流量特征的差异性量化为数学距离,进而对 6 种流量实施分类,样本精确率达到了 90% 以上。文献^[51]提出了一种基于 SVM (Support Vector Machine) 的流量分类算法,通过构建超平面进行流量分类,具有分类程度高、泛化性高和计算快速的优点。此外,还有基于决策树和神经网络的分类算法^[52-54]。

在网络异常检测方面,主要有基于统计分析、机器学习、数据挖掘以及小波

分析的异常检测。基于统计分析的检测方法^[55-56]是异常检测的最基本的方法，通过采样收集用户数据，利用样本的特征参数值描述用户行为，进一步与用户历史行为进行对比，检查当前用户行为特征与历史特征的偏差是否在允许范围内以确定异常情况。基于机器学习的异常检测^[57]同样是通过现有数据与历史数据的对比，但更注重异常检测系统的自我学习和自我完善过程。Lee 等人将数据挖掘方法首次应用到了网络流量异常检测中^[58]，通过建立正常网络流量模式对比发现网络流量的异常现象。Alarcon 等人^[59]2002 年提出了利用小波分析进行网络异常检测，其中 Haar 小波是最先提出并得到广泛使用的时域不连续小波函数。以上研究方法均通过不同的数学方法分析业务流在网络异常环境下的特征属性，进而预测未知的网络事件。

1.2.3 已有研究存在的不足

网络业务流行为复杂性的根源在于海量业务共存于网络中，由于网络的服务能力有限造成大量业务相互耦合，使得业务间的关联关系更为紧密，表现出的宏观现象更难以理解。通过总结近年来网络业务流行为研究现状，不难看出目前网络业务流行为分析方法大多集中在业务流的生成、到达、以及业务分布间隔的规律上，利用这些参数分析网络业务行为一方面很难与网络用户息息相关的业务服务质量直接关联；另一方面，虽然已有研究考虑到业务行为时间上或空间上的关联性，但对于业务服务质量的耦合关联性极少涉及。无论从网络自身还是用户体验的角度，挖掘业务流服务质量间的相互影响规律可以更好地理解网络业务行为，高效提升网络整体性能。此外，以往研究成果大多基于稳定网络环境，而真实网络瞬息万变，网络结构和业务分布均可能在短时间内发生巨大变化，因此，动态网络环境中业务流行为特性也是本文研究的重点。

综上，目前已有研究主要存在以下不足：

- （1）已有研究大多分析业务的宏观行为属性，很少考虑到业务服务质量之间的微观耦合关系，微观关系的建立更能深入解释网络复杂现象产生的原因；
- （2）已有研究未考虑复杂动态网络环境中业务的行为特征，例如各种复杂网络结构、突发网络流量、网络故障环境等；
- （3）已有研究对网络业务行为的动力学特征尚未研究充分，例如网络业务流行为的动态传播特性等。

1.3 研究内容及关键问题

本文致力于挖掘业务流服务质量间的耦合关联关系,综合考虑网络拓扑、网络资源分配、路由策略、通信协议等因素对业务流行为模式的影响,建立网络业务流服务质量耦合关联模型,并基于此分析网络突发和故障环境下的业务流传播动力学行为,进而提出有效的传播控制方案。

(1) 基于复杂网络的业务流服务质量耦合行为建模与机理研究

由于节点队列缓存长度与链路带宽容量有限,导致流经同一节点或同一链路的业务之间对于网络资源的竞争十分激烈,业务流之间呈现明显的耦合相关性;同时,每条业务经过网络中很多的节点与链路,在这些网络单元中,又与其它节点和链路承载的业务进行耦合关联。全网业务通过节点与链路紧密耦合,呈现密切关联特性,使得大规模网络中业务间的耦合机理分析极为困难。本文拟利用网络系统与弹性物理网络系统的相似性,基于弹性力学相关理论,建立业务服务质量弹性物理耦合模型,并通过将业务间的耦合作用扩展至整个网络,构建业务流耦合关联网络:节点表示业务流,连边表示业务服务质量间耦合关联程度。基于上述模型,研究网络中业务流服务质量间的耦合机理以及业务关联关系体现出的复杂网络特性。

(2) 基于业务流突变的服务质量波动传播特性与控制算法研究

网络状态和业务需求往往是动态、随机变化的,任何业务带宽的突变都会对网络中其他业务流的服务质量产生影响。由于网络资源有限,业务带宽的突增将对与其共享网络资源的其他业务产生挤压作用,使其带宽减小,并将此作用通过自身以及受影响的其他业务逐级扩散至全网。带宽波动影响将涉及到的网络中大面积的网元和业务流,如何分析和刻画业务流间的服务质量波动传播规律成为当前亟待解决的问题。本文拟基于业务流耦合关联网络,研究突变业务流对全网业务服务质量波动影响效应,具体包括不同网络环境下突变流影响范围和传播强度的变化,并分析造成波动影响的内在原因。此外,拟利用业务的社团属性,并结合传染病模型中“免疫”的概念,研究基于社团结构的免疫控制方案,有效抑制突变业务流对全网业务的扩散影响。

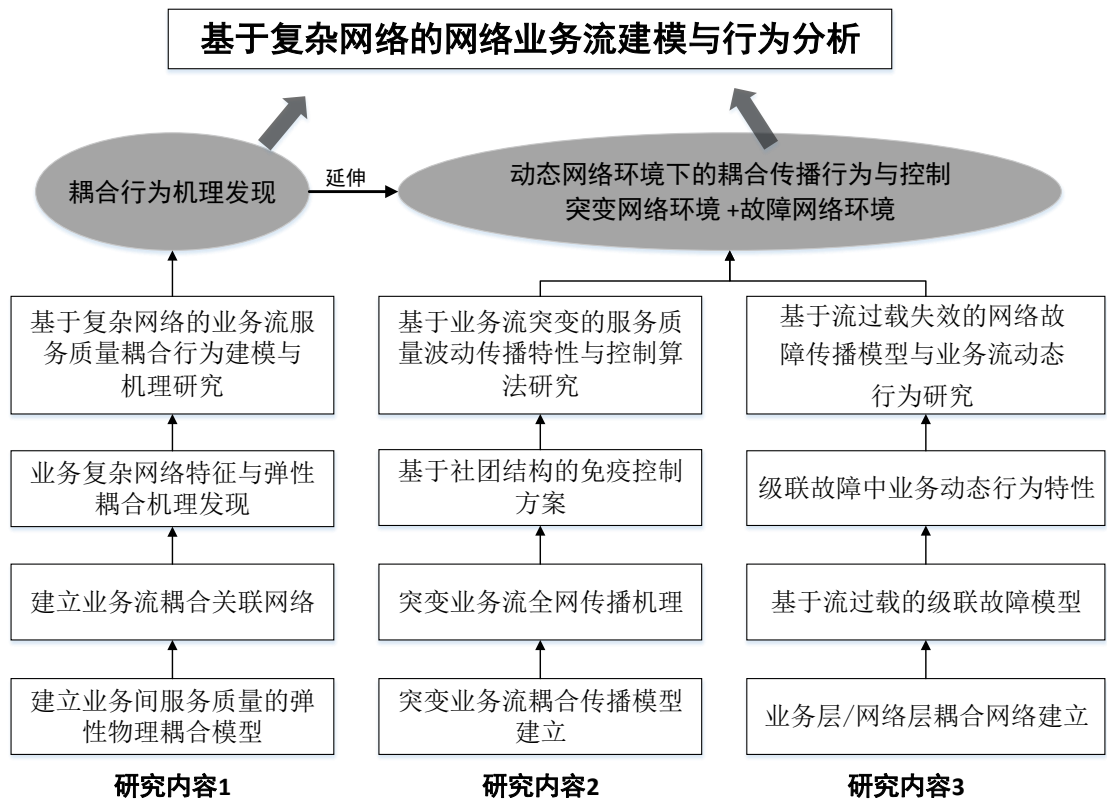


图 1-4 研究内容

(3) 基于流过载失效的网络故障传播模型与业务流动态行为研究

设备故障、自然灾害、网络恶意攻击都会引起的网元之间连接关系的变化，并对业务的选路、资源分配等产生重要影响。因此，网络中的耦合行为不仅体现在业务与业务之间，也体现在物理链路、节点与被传送的业务之间。网络连接关系的变化，将导致业务传送中断、网络资源重新分配。由于网络传送的业务数目极为庞大，网络节点的动态故障会引起业务的大规模动态迁移，业务此时的动态变化规律对于保障网络服务质量、提升网络安全可靠性具有指导意义。本文拟基于节点、链路与业务间的支撑依赖关系，建立网络层与业务层的双层网络，分析流量过载引发的网络故障级联效应，并探索级联效应中业务流动态行为变化规律，如传输行为、竞争行为等。

图 1-4 显示了本文研究内容之间的内在关联。研究内容 1 从发现业务流耦合行为机理入手，为研究内容 2 和 3 的开展奠定基础。研究内容 2 和 3 分别从突变网络环境和故障网络环境两个角度共同探索业务流的传播动力学行为。三项内容共同支撑了本文的研究工作。

1.4 论文组织安排

本文后续章节安排如下：

第二章介绍与本文研究内容相关的复杂网络基础知识，包括复杂网络度量参数、典型网络演化模型等。第三章主要研究复杂网络中业务流服务质量间的耦合行为。通过构建业务流弹性物理模型以及业务流耦合关联网络，对业务服务质量间的关联行为进行探究。第四章主要研究突变流对业务服务质量的波动传播特性及控制算法。利用突变业务流和其他业务共享资源的直接或间接关系，建立突变业务流的服务质量影响传播模型，并依据业务流的社团特性，提出基于社团结构的突变业务流免疫策略。第五章主要研究基于流过载的网络故障传播与业务流动态行为。通过构建网络层与业务层的双层网络，提出基于链路过载失效的级联故障模型，考察链路级联失效过程中业务的动态行为变化。第六章是总结和展望，主要总结本文的研究成果、创新点以及今后的研究方向。最后是附录、致谢和发表论文部分。

1.5 本章小结

本章首先介绍了论文的研究背景，分析了当前网络业务流行为研究面临的新挑战；其次，总结了网络业务流行为的已有研究成果与应用现状，并分析了当前研究的不足之处；再次，介绍了本文的主要研究内容及关键问题；最后，介绍了论文的组织结构安排，为后续章节的展开做铺垫。

参考文献

- [1] Perelló, J., Spadaro, S., Ricciardi, S., et al. All-optical packet/circuit switching-based data center network for enhanced scalability, latency, and throughput[J]. IEEE Network, 2013, 27(6):14-22.
- [2] Mathison, S. L., Roberts, L. G., Walker, P. M. The history of telenet and the commercialization of packet switching in the U.S[J]. IEEE Communications Magazine, 2012, 50(5):28-45.
- [3] Furukawa, H., Mendinueta, J. M. D., Wada, N., et al. Spatial and spectral super-channel optical packet switching system for multigranular SDM-WDM

- optical networks[J]. IEEE/OSA Journal of Optical Communications & Networking, 2017, 9(1):A77-A84.
- [4] 运行监测协调局. 2018 年 2 月份通信业经济运行情况[EB].
<http://www.miit.gov.cn/newweb/n1146312/n1146904/n1648372/c6104255/content.html>, 2018-3-23.
- [5] Dahlman, E., Parkvall, S., Skold, J. 4G: LTE/LTE-Advanced for Mobile Broadband[M]. 东南大学出版社, 2012:iv.
- [6] Tu, G., Li, C., Peng, C., et al. VoLTE:A Lightweight Voice Solution to 4G LTE Networks[A].// International Workshop on Mobile Computing Systems and Applications [C]. New York: ACM, 2016:3-8.
- [7] Toukabri, T., Said, A. M., Abd-Elrahman, E., et al. Distributed D2D Architecture for ITS Services in Advanced 4G Networks[A].// Vehicular Technology Conference[C]. Boston: IEEE, 2016:1-7.
- [8] Huang, J., Qian, F., Gerber, A., et al. A close examination of performance and power characteristics of 4G LTE networks[A].// International Conference on Mobile Systems, Applications, and Services[C]. Long Wood Bay: ACM, 2012:225-238.
- [9] Garc ía, M., Daly, A. J., S á nchez-Cabezudo, S. Identifying the new Influences in the Internet Era: Social Media and Social Network Analysis[J]. Rev. Esp. Investig. Sociol., 2016, 153: 23-40.
- [10] Kristoufek, L. BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era[J]. Scientific Reports, 2013, 3(3):3415.
- [11] Zhao, H., Li, L., Peng, H., et al. Fixed-time synchronization of multi-links complex network[J]. Modern Physics Letters B, 2017, 31(02):1-87.
- [12] Rathkopf, C. Network representation and complex systems[J]. Synthese, 2018, 195(1):55-78.
- [13] Liu, C., Arunkumar, N. Risk prediction and evaluation of transnational transmission of financial crisis based on complex network[J]. Cluster Computing,

2018, 10:1-7.

- [14] 汪小帆, 李翔, 陈关荣. 网络科学导论[M]. 高等教育出版社, 2012.
- [15] Garcia-Luna-Aceves, J. J., Sevilla, S., et al. A Simple Solution to Scale-Free Internet Host Mobility[A]// International Conference on Computer Communication and Networks[C]. USA: IEEE, 2017:1-9.
- [16] Chirgui, Z. M. Small-world or Scale-Free Phenomena in Internet: What Implications for the Next-generation Networks? [J]. Review of European Studies, 2012, 4(1): 85-93.
- [17] Hayashi, Y. Tomographies of Internet and WWW: Sampling of the Scale-Free Networks[J]. Journal of the Institute of Electronics Information & Communication Engineers, 2006, 89(4):322-325.
- [18] Barabási, A. L., Albert, R., Jeong, H. Scale-free characteristics of random networks: the topology of the world-wide web[J]. Physica A Statistical Mechanics & Its Applications, 2012, 281(1):69-77.
- [19] Sohn, I. Small-World and Scale-Free Network Models for IoT Systems[J]. Mobile Information Systems, 2017, 2017(61):1-9.
- [20] Jarman, N., Steur, E., Trengove, C., et al. Self-organisation of small-world networks by adaptive rewiring in response to graph diffusion[J]. Scientific Reports, 2017, 7(1):13158.
- [21] 运行监测协调局. 2017 年通信业统计公报[EB].
<http://www.miit.gov.cn/n1146290/n1146402/n1146455/c6048084/content.html>, 2018-2-2.
- [22] Jin, X., Min, G. Modelling and Analysis of an Integrated Scheduling Scheme with Heterogeneous LRD and SRD Traffic[J]. IEEE Transactions on Wireless Communications, 2013, 12(6):2598-2607.
- [23] Muscariello, L., Mellia, M., Meo, M., et al. Markov models of internet traffic and a new hierarchical MMPP model [J]. Computer Communications, 2005, 28(16):1835-1851.
- [24] Fasen, V. Modeling network traffic by a cluster Poisson input process with heavy

- and light-tailed file sizes[M]. J. C. Baltzer AG, Science Publishers, 2010.
- [25] Martin, R. D., Yohai, V. J. 4 Robustness in time series and estimating ARMA models[J]. Handbook of Statistics, 2016, 5(85):119-155.
- [26] Han, Y., Jeong, T., Yoo, J. H., et al. FLAME: Flow level traffic matrix estimation using poisson shot-noise process for SDN[A].// Netsoft Conference and Workshops[C]. Seoul: IEEE, 2016:102-106.
- [27] Ko, E., Ahn, J., Kim, E. Y. 3D Markov Process for Traffic Flow Prediction in Real-Time[J]. Sensors, 2016, 16(2):147.1-147.19.
- [28] 李文璟, 陈晨, 喻鹏,等. 基于 S-ARIMA 模型的无线通信网络业务量预测方法[J]. 北京邮电大学学报, 2017, 40(S1):10-14.
- [29] Wen, Y., Zhu, G., Xie, C., et al. Fractional Autoregressive Prediction for Long Range Bursty Traffic[J]. Computer Science, 2009, 36(7):79-81.
- [30] Le, L. Long-range dependence analysis of Internet traffic[J]. Journal of Applied Statistics, 2011, 38(7):1407-1433.
- [31] Matrawy, A., Lambadaris, I., Huang, C. MPEG4 traffic modeling using the transform expand sample methodology[A].// Proceedings of International Workshop on Networked Appliances[C]. Gaithersburg: IEEE, 2002:249-256.
- [32] Bregni, S. Compared Accuracy Evaluation of Estimators of Traffic Long-Range Dependence[J]. IEEE Latin America Transactions, 2016, 13(11):3649-3654.
- [33] Leland, W. E., Taqqu, M. S., Willinger, W., et al. On the self-similar nature of Ethernet traffic[A].// ACM SIGCOMM Computer Communication Review[C]. San Francisco: ACM, 1993:183-193.
- [34] Pramanik, S., Datta, R., Chatterjee, P. Self-similarity of data traffic in a Delay Tolerant Network[A].// Porto: Wireless Days[C]. IEEE, 2017.
- [35] Porshnev, S. V., Bozhalkin, D. A. The study of self-similarity of the traffic transmitted in the backbone Internet channel[A].// Dynamics of Systems, Mechanisms and Machines[C]. Omsk: IEEE, 2017:1-7.
- [36] Benhaddou, R., Datta, S., Koul, H. L. Deconvolution model with fractional

- Gaussian noise: A minimax study[J]. *Statistics & Probability Letters*, 2016, 117:201-208.
- [37] Artemov, A. V., Burnaev, E. V. Optimal estimation of a signal, observed in a fractional Gaussian noise[J]. *Theory of Probability & Its Applications*, 2015, 60:163–171.
- [38] Delignières, D. Correlation Properties of (Discrete) Fractional Gaussian Noise and Fractional Brownian Motion[J]. *Mathematical Problems in Engineering*, 2015(2015): 485623.
- [39] Bondarenko, V. The Model of Real Data Constructing Using Fractional Brownian Motion[J]. *International Journal of Swarm Intelligence & Evolutionary Computation*, 2017, 06(1):1000149.
- [40] Sun, X., Li, Y., Zhang, M. Predicting Bursty Network Traffic with Self-Similarity Characteristic over Echo State Covariation Orthogonality Network[J]. *International Journal of Multimedia & Ubiquitous Engineering*, 2017, 12(5):29-42.
- [41] Matulin, M., Štefica, M., Adlešić, D. Application of the M/G/1/ ∞ / ∞ /FCFS Queueing System for the Internet Node Performance Analysis[A].// *Research Conference in Technical Disciplines*[C]. Slovačka: RCITD, 2016:15-18.
- [42] Shen, L., Zhang, Y., Luan, J., et al. An Energy Correction Wavelet Model for traffic in terminal communication access network[A].// *International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*[C]. Changsha: IEEE, 2016:1986-1990.
- [43] Crovella, M. E., Taqqu, M. S., Bestavros, A. Heavy-Tailed Probability Distributions in the World Wide Web[J]. *A Practical Guide To Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed distributions*, 1998:3-25.
- [44] Downey, A. B. Lognormal and Pareto distributions in the Internet[J]. *Computer Communications*, 2005, 28(7):790-801.
- [45] Sarvotham, S., Riedi, R., Baraniuk, R. Network and user driven alpha-beta on-off source model for network traffic[J]. *Computer Networks*, 2005, 48(3):335-350.

- [46] Brownlee, N., Claffy, K. C. Understanding Internet traffic streams: dragonflies and tortoises[M]. IEEE Press, 2002.
- [47] Lan, K. C., Heidemann J. A measurement study of correlations of Internet flow characteristics[J]. Computer Networks, 2006, 50(1):46-62.
- [48] Park, C., Hernández-Campos, F., Marron, J. S., et al. Analysis of dependence among size, rate and duration in internet flows[J]. Annals of Applied Statistics, 2010, 4(1):26-52.
- [49] Moore, A.W., Zuev, D. Internet traffic classification using bayesian analysis techniques[A].// ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems[C]. Banff: ACM, 2005:50-60.
- [50] Huang, S., Chen, K., Liu, C., et al. A statistical-feature-based approach to internet traffic classification using Machine Learning[A].// International Conference on Ultra Modern Telecommunications & Workshops[C]. St. Petersburg: IEEE, 2009:1-6.
- [51] Li, X., Qi, F., Xu, D., et al. An Internet Traffic Classification Method Based on Semi-Supervised Support Vector Machine[A].// IEEE International Conference on Communications[C]. Kyoto: IEEE, 2011:1-5.
- [52] Tong, D., Qu, Y. R., Prasanna, V. K. Accelerating Decision Tree Based Traffic Classification on FPGA and Multicore Platforms[J]. IEEE Transactions on Parallel & Distributed Systems, 2017, 99:1-1.
- [53] Su, L., Mu, H., Min, Y., et al. The research of classified method of the network traffic in security access platform based on decision tree[A].// IEEE International Conference on Software Engineering and Service Science[C]. Beijing: IEEE, 2017:475-480.
- [54] Wang, W., Zhu, M., Zeng, X., et al. Malware traffic classification using convolutional neural network for representation learning[A].// International Conference on Information Networking[C]. Da Nang: IEEE, 2017:712-717.
- [55] Ristic, B., Scala, B. L., Morelande, M., et al. Statistical analysis of motion patterns in AIS Data: Anomaly detection and motion prediction[A].//

- International Conference on Information Fusion[C]. Cologne: IEEE, 2008:1-7.
- [56] Moustafa, N., Slay, J. The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set[J]. Information Systems Security, 2016, 25(1-3):18-31.
- [57] Shon, T., Moon, J. A hybrid machine learning approach to network anomaly detection[J]. Information Sciences, 2007, 177(18):3799-3821.
- [58] Lee, W., Nimbalkar, R. A., Yee, K. K., et al. A Data Mining and CIDF Based Approach for Detecting Novel and Distributed Intrusions[J]. Lecture Notes in Computer Science, 2000, 1907:49-65.
- [59] Alarcon-Aquino, V., Barria, J. A. Anomaly detection in communication networks using wavelets[J]. IEE Proceedings - Communications, 2002, 148(6):355-362.

第二章 复杂网络相关理论

无论是关系到人们基本生活的通信网络、电力网络、交通网络，还是深入到社会各方面的经济网络、政治网络、社交网络，人们已经生活在一个复杂交织的网络世界中。近十几年来，复杂网络理论发展迅速，成为描述复杂系统极其重要且十分有效的工具，相关研究遍及社会学^[1-3]、生物学^[4-6]、计算机科学^[7-11]等方面。本文在建立模型和分析问题时利用了大量复杂网络相关知识，为后文能够方便地进行引用，本章简要地介绍复杂网络相关基础理论。

2.1 网络及其表示方法

所谓网络，是由多个节点及节点之间的连边组成的集合。网络通常用 $G(V, E)$ 表示， V 表示节点集合， E 表示连边集合。网络连边 e 由节点对 (u, v) 表示。如果任意节点对 (u, v) 和 (v, u) 表示同一条连边，那么连边 e 为无向边，对应的网络称为无向网络；否则，为有向网络。如图 2-1 所示，(a) 为无向网络，(b) 为有向网络。

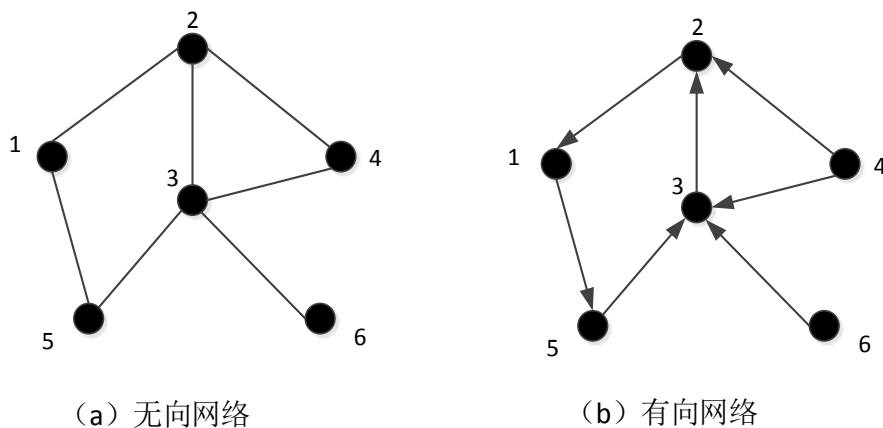


图 2-1 无向网络与有向网络

在数学上，利用邻接矩阵 A (Adjacency Matrix) 可以简单地表示网络。邻接矩阵 A 中元素 a_{ij} 的含义如下：

$$a_{ij} = \begin{cases} 1, & \text{如果节点 } i \text{ 和节点 } j \text{ 之间存在一条边} \\ 0, & \text{其他} \end{cases} \quad (2-1)$$

例如，图 2-1 (a) 无向网络的邻接矩阵可以表示为公式 (2-2)，(b) 有向网络的邻接矩阵可以表示为公式 (2-3)。

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (2-2)$$

$$A' = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (2-3)$$

无向网络的邻接矩阵为对称矩阵，有向网络的邻接矩阵为非对称矩阵。如果网络中所有连边的长度均为 1，即对于 $\forall(i, j) \in E$ ， $a_{ij} = 1$ ，则称该网络为无权网络。在一些情况下，给边赋予强度或权重等含义非常有必要，称之为加权网络，可通过将邻接矩阵中对应元素的值设定为相关权重值进行表示。对于 Internet，边的权重可以表示节点间的数据流量或带宽。邻接矩阵可以反映网络的拓扑信息，是研究复杂网络结构特征的重要表示工具。

2.2 复杂网络特征度量

度量复杂网络结构特性的方法包括度分布、平均路径长度、聚集系数、介数、模块度等。

度分布 (Degree Distribution)。 网络节点度可以利用邻接矩阵表示为：

$$k_i = \sum_{j=1}^N a_{ij} \quad (2-4)$$

其中， k_i 表示节点*i*的度数，即与该节点连接的所有边的数目总和。

在无向网络中，任一边都有两个端点，如果网络共有 m 条边，那么网络节点度和边存在以下关系：

$$2m = \sum_{i=1}^N k_i \quad (2-5)$$

有向网络中节点度的计算较为复杂。入度（In-degree）是连接到该节点的入边数目，出度（Out-degree）是出边数目。用邻接矩阵将其分别表示为：

$$\begin{cases} k_i^{out} = \sum_{j=1}^N a_{ij} \\ k_i^{in} = \sum_{j=1}^N a_{ji} \end{cases} \quad (2-6)$$

度分布是网络的一个重要统计特征，是对网络中节点度数的总体描述，即节点度的概率分布。度分布 $P(k)$ 是节点有 k 条连边的概率。在目前的研究中，较为常见的有泊松分布和幂律分布。随机网络的度分布满足泊松分布。泊松分布的特点是当 k 很小或很大的时候， $P(k)$ 值都接近零；而在某一特定值处将达到峰值，即大多数节点的度数都分布在这一特定值附近。无标度网络的度分布为幂律分布，即 $P(k) = k^{-\gamma}$ ， γ 为常数。 γ 不同，网络表现出的特性也不同。

平均路径长度（Average Path Length）。也称作特征路径长度。作为衡量网络结构特性的基本度量参数，能够反映网络在信息传递方面性能和效率的高低，表达式为：

$$APL = \frac{1}{N(N-1)} \sum_{\substack{i, j \in V \\ i \neq j}} d_{ij} \quad (2-7)$$

其中， N 表示网络节点数目， d_{ij} 表示节点*i*和节点*j*之间的最短路径长度。本文使用 Dijkstra 最短路径算法进行业务选路。

聚集系数（Clustering Coefficient）。节点聚集系数体现了一个节点的邻居节点的聚合程度。在实际网络中，节点偏向于与具有相似属性的节点建立关联。例如社交网络中朋友圈中的成员往往互相熟知，科学引文网络中相似研究方向的科研成果往往相互引用。节点之间的紧密关联性可以利用聚集系数量化表示，在无

向网络中，节点 i 的聚集系数表示为：

$$C_i = \frac{n}{C_{k_i}^2} = \frac{2n}{k_i(k_i - 1)} \quad (2-8)$$

其中， k_i 是节点 i 的度数； n 表示节点 i 的邻居之间存在的连边数； $C_{k_i}^2$ 表示邻居间可能存在的最大连边数目。将网络中所有节点的聚集系数求平均，可得到网络的聚集系数：

$$C = \frac{1}{N} \sum_{i \in N} C_i \quad (2-9)$$

其中， N 为网络节点总数。

介数 (Betweenness)。介数的思想由 Freeman 1977 年提出^[12]。作为全局变量，它反映了节点或边在网络连通能力中的影响力。介数分为点介数和边介数，定义为：

$$\text{点/边介数} = \frac{\text{经过某节点/边的最短路径数目}}{\text{网络中所有最短路径数目}} \quad (2-10)$$

具有高介数的节点和连边控制着其他节点间消息的传递，在网络中有很强的影响力。在消息传递中，高介数的节点流经的消息最多，如果对节点进行监控，将高效率地把控网络大范围的信息；而删除高介数节点或边，将破坏网络的连通效率。虽然在现实网络中信息不一定沿着最短路径进行传递，但介数中心性仍然对信息的流动有着至关重要的影响。

模块度 (Modularity)。模块也称作社区或社团 (Community)。网络的社团特性有两个显著特征：一是社团内部节点的连接非常密集；二是社团与社团之间只有少数甚至没有连接，即社团之间有清晰的划分边界。网络社团结构的划分有助于理解与分析网络的组织特性。例如在科学家网络中，同一社团的学者一般有共同的研究领域；在电影演员网络中，曾经合作过或主演相同题材的演员大多在同一社团。模块度的定义由 Girvan 和 Newman 在 2004 年提出^[13]，公式表示为：

$$Q = \sum_i e_{ii} - a_i^2 \quad (2-11)$$

其中， $a_i = \sum_j e_{ij}$ 指连接社团 i 中连边的比例， e_{ii} 表示两个端点都在社团 i 中边的

比例。模块度越大，社团结构越明显。

最大连通片（Giant Component）。连通片也称作连通分支或分图，连通片中任何两个节点之间均存在路径。网络一般由多个连通片组成，其中最重要的是最大连通片，最大意味着该分支的节点数目最多，而任何其它节点或边的加入都将使分支的连通性遭受破坏。通常，最大连通片的大小能反映网络整体可靠性的高低。公式表示为：

$$g = \frac{N'}{N} \quad (2-12)$$

其中， N' 为最大连通片的节点数目， N 为网络总节点数。当 $N \rightarrow \infty$ 时， g 成为衡量网络遭受攻击时鲁棒性高低的一个重要参数，对于不同类型网络该比值会呈现不同的结果。

2.3 典型复杂网络结构

（1）规则网络（Regular Network）

规则网络是指每个节点的邻居个数都相同的网络，即每个节点的度数都相同。这里介绍一种经典的规则网络：最近邻耦合网络。“最邻近”是指节点仅与其距离最近的其他节点建立连接，算法过程如图 2-2 所示。

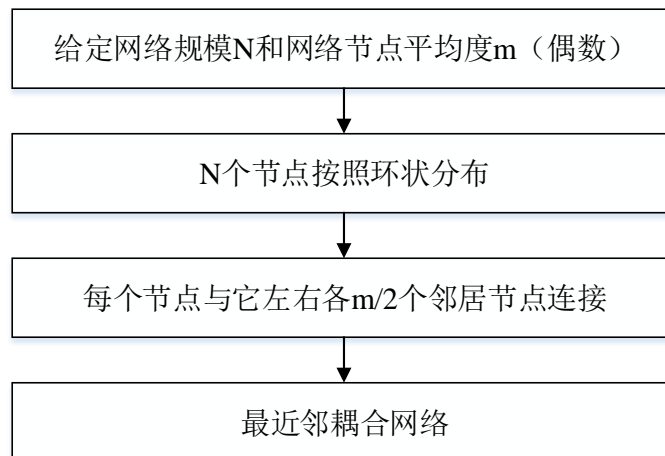


图 2-2 最近邻耦合网络

（2）随机网络（Random Network）

随机网络中任两个节点之间的连边都是依据某一概率产生的，虽然每次生成的网络都不同，但在结构测度方面有一定的规律性，如最短路径长度、度分布等。最经典的随机网络模型由 Erdős 和 Rényi 在 1959 年提出^[14]，称为 ER 随机图模型。算法过程如图 2-3 所示。

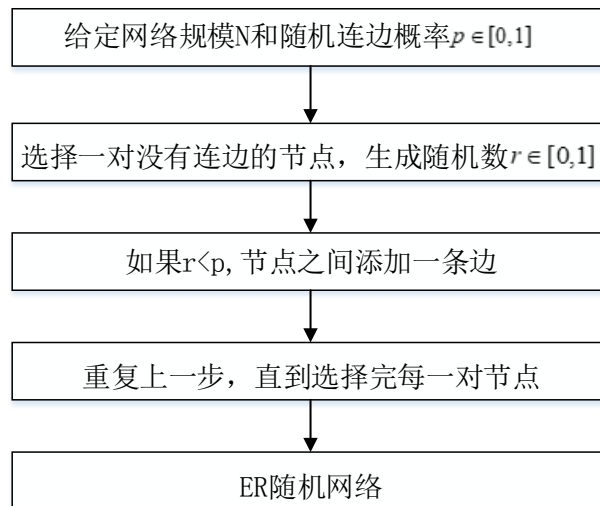


图 2-3 ER 随机网络

（3）小世界网络（Small World Network）

判定一个网络是否为小世界网络一般使用两个特征值衡量：特征路径长度和节点聚集系数。小世界网络的特征路径长度通常较小，一般稍大于同等规模的随机网络；聚集系数却非常高，通常远远高于随机网络。

规则网络具有很高的聚集系数，少量“捷径”的建立能够迅速减少路径长度，但对聚集系数的影响却不大。基于该思想，Watts 和 Strogatz 在 1998 年提出了 WS 小世界网络模型^[15]，算法过程如图 2-4 所示。

WS 模型随着“捷径”重连概率 p 的变化表现出不同的网络结构特点。以图 2-5 为例， $p=0$ 时，模型得到规则网络，网络平均节点度为 4； $0 < p < 1$ 时，模型得到 WS 小世界网络，相比于规则网络，增加了节点间的“长程连接”； $p=1$ 时，模型得到随机网络，网络连边随机产生。可以看出， p 值的改变，可以使网络完成由规则网络向随机网络的转变。

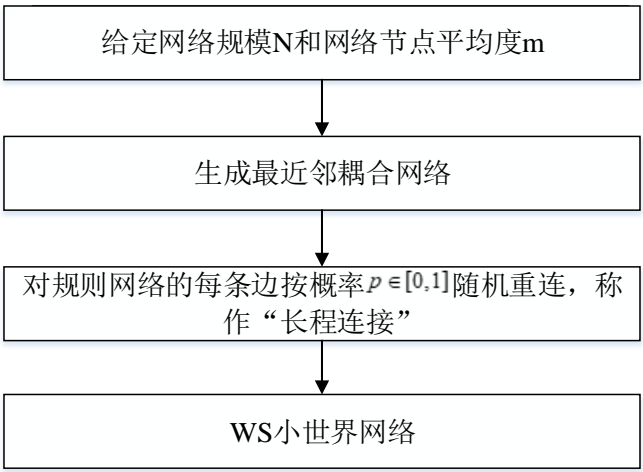


图 2-4 WS 小世界网络

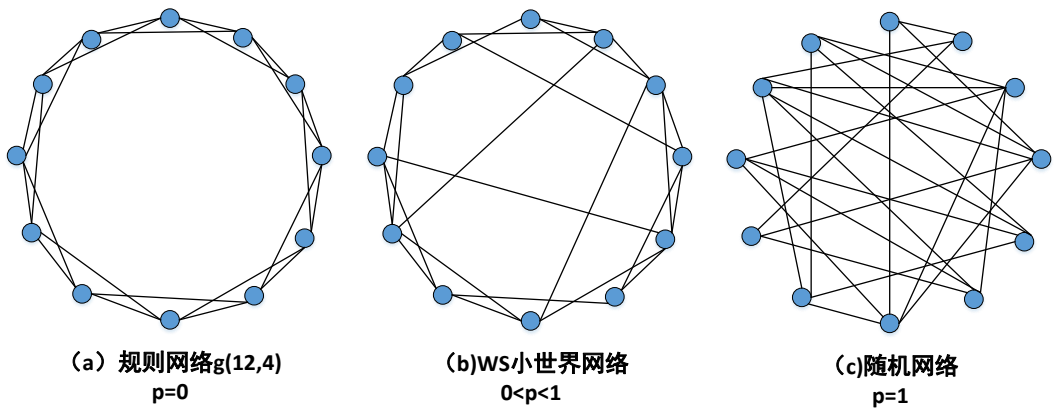
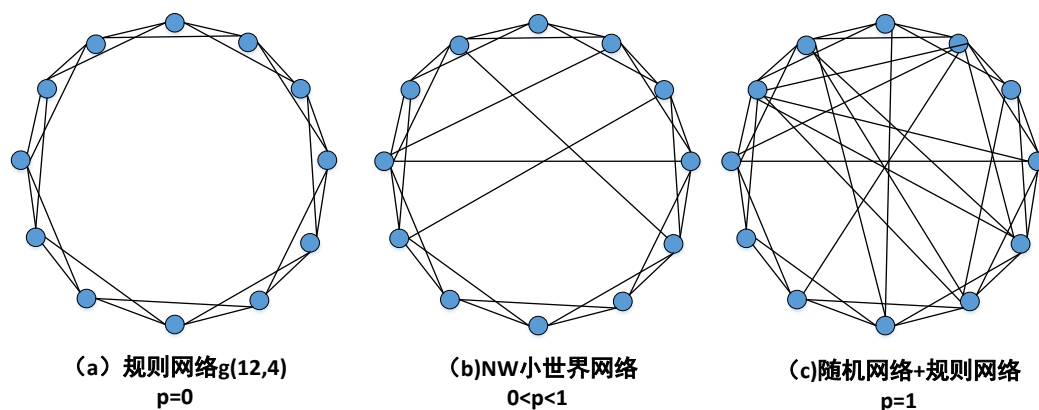


图 2-5 WS 模型随机重连过程^[15]

WS 模型需要断掉连边重新连接，易产生孤立簇，对后续网络结构的研究造成困难。Newman 和 Watts 针对 WS 模型的缺陷进行了改进，提出了 NW 改进模型^[16]。模型摒弃了 WS 模型的“随机重连”原则，改为保持初始规则网络连边不变，依概率 p 增加连边作为长程连接，增加连边相比于更改连边避免了网络孤立簇的产生。

NW 模型的“随机加边”过程如图 2-6 所示。与 WS 模型随概率 p 变化的结果不同的是， $p=1$ 时，模型得到规则网络和随机网络的叠加结果，而非完全的随机网络。

图 2-6 NW 模型随机加边过程^[16]

(4) 无标度网络 (Scale Free Network)

事实上，ER 模型和 WS 模型并不能完全表示许多真实网络的结构特性，具有一定的局限性。因此，为更好地描述现实网络，1999 年，Barabási 和 Albert^[17]通过大量研究证实了许多真实世界网络的度分布都具有幂律特性，并称之为无标度网络。幂律分布体现了网络的高自组织性，大多数节点更倾向于连接到度数更高的“枢纽”节点。网络中少数的“枢纽”节点对网络的运行有至关重要的作用，尤其是对网络最短路径长度的影响。网络连边增长和择优连接是生成无标度网络的两条必要机制。具体而言，算法过程如图 2-7 所示。

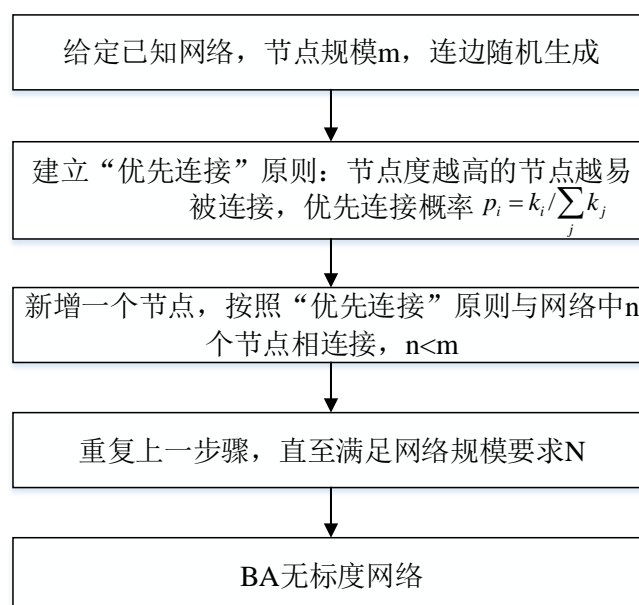


图 2-7 BA 无标度网络

2.4 通信领域复杂网络研究进展

通信领域复杂网络的研究重点目前主要有网络结构、网络形成机制、网络演化等；同时，许多学者在网络稳定性、鲁棒性等方面也有深入研究。通信网络复杂属性的研究有助于深入理解网络业务流行为表征的内在机理。

在网络结构研究方面，大致经历了三个阶段。第一阶段：学者们认为网络节点间的关系可以用规则网络表示，例如最近邻耦合网络。第二阶段：1959 年，Erdős 和 Rényi^[14]提出了一种新型结构的 ER 随机网络模型，认为节点之间的连边不再唯一确定，而是依据概率决定，ER 随机网络在近 40 年的时间里被认为是最符合实际网络特点模型。第三阶段：直到 20 世纪 90 年代末，一些学者发现许多真实网络并不完全满足规则网络或随机网络的特性，而是具有新的网络特征，其中最具代表性的是小世界网络和无标度网络。

小世界网络模型由 Watts 和 Strogatz^[15]于 1998 年提出，小世界网络一般具有较短的特征路径长度和极高的聚集系数。此后，Newman 和 Watts^[16]对 WS 模型进行了改进，提出了 NW 小世界网络模型，有效避免了网络孤立簇的产生。无标度网络特性由 Barabási 和 Albert^[17]在探究 WWW 网络的动态演化过程中发现。WWW 网络具有大规模的自组织特性，网络中的节点更趋向于连接具有更高度数的节点，网络的节点度服从幂律分布 $P(k) \sim k^{-\gamma}$ ： k 是节点度， γ 是一个常数。网络小世界特征和无标度特征的发现，推动复杂网络进入全新阶段，科学家发现包括因特网、万维网、传感器网在内的许多现实网络都可以由复杂网络的无标度特性或小世界特性进行描述。以因特网为例，Faloutsos 和 Yook^[18-19]等人发现 Internet 的聚集系数高于相同规模的随机网络，同时特征路径长度较短，约在 3.70~3.77 之间，满足小世界属性；度分布也符合幂律分布，满足无标度属性。Zhou S 等人发现了 Internet 节点的 rich-club 现象，即极少数的节点却有着极大数目的连接度数，同样也符合幂律分布的异质特性^[20]。Internet 复杂属性的发现对于科学家深入研究 Internet 的演化机制和结构特性具有重要意义。

在网络演化模型方面，主要有静态模型和动态模型两类。动态模型能够反映网络的节点和连边的增长特征，体现网络形成的动态机理，更符合网络随时间增长的动态过程。目前，比较常见的 Internet 演化模型有 BA (Barabási-Albert) 模型，GLP (Generalized Linear Preference) 模型，IG (Interactive Growth) 模型，HOT (Highly Optimized Tolerance) 模型等。Barabási 和 Albert^[17]提出的 BA 模型

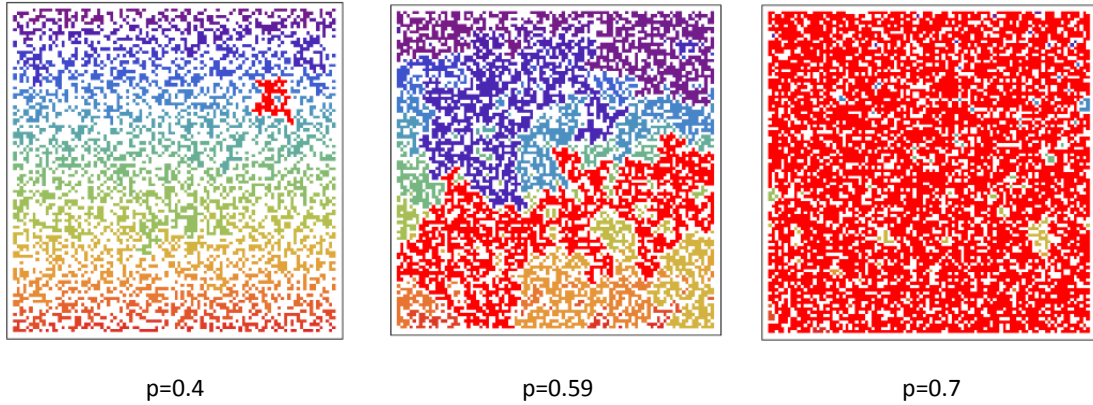


图 2-8 渗流过程^[33]。图为 100*100 的网格网络，最大连通片用红色进行标识。

主要体现网络的增长特性和优先连接特性，新加入节点将按照某种规则优先与已有节点进行连接。GLP 模型由 Tian Bu 等人^[21]提出，该模型同时体现了度的幂律特征和网络结构的“小世界”特性。由于 BA、GLP 模型中节点和连边的增加过程是相互独立的，2003 年 Zhou 提出了节点和边联合增长模式的 IG 模型^[22]，该模型体现出的 rich-club 特性更接近于真实网络情况。以上模型的演化原理大多基于新增节点的“优先连接”特性，并未考虑到节点和连边的重新调整，不完全符合 Internet 的实际情况。HOT 模型是一种多目标优化的拓扑演化模型^[23]，综合考虑收益、资源成本和风险承受这三者之间的平衡状态，合理评估新的节点和连边带来的增益，及时进行结构调整，是一种高度优化设计的结果。

在网络稳定性和鲁棒性方面，“渗流理论”^[24]成为研究复杂网络鲁棒性的基本方法。其问题描述如下：假设网络中有 $1-p$ 比例的节点失效（人为攻击或自然灾害），与失效节点相连的链路也随之失效，网络被分割成若干连通片以及孤立的节点。通常假设只有最大连通片中的节点和链路具有功能，最大连通片所占的比例 P_∞ 则是衡量网络鲁棒性的指标。图 2-8 展示了 100*100 的网格网络随 p 值变化的渗流过程。图中 p 值的增加过程同时也是失效节点 $1-p$ 减少的过程，网络最大连通片的比例在此过程中显著增大。 P_∞ 是关于 p 的不减函数，且函数曲线可能存在相变临界值 p_c 。 p_c 值越小，网络鲁棒性越高。研究表明，ER 网络对于随机攻击鲁棒性很差，对蓄意攻击鲁棒性很强，而 BA 网络对随机攻击却有很强的鲁棒性，对蓄意攻击鲁棒性很差^[25-26]。级联失效中另一个较为经典的情形是“流的过载失效”，例如互联网中业务流分配不均衡导致的节点/链路级联失效。研究表明当节点承载的业务流负荷超过了其最大承受能力时将会发生失效，而失效的节点会导致网络负载重新分配，进而使得其他节点失效，最后发生网络的彻

底崩溃。相关现象在电力网络，交通网中也有明显的体现^[27-29]。此外，一些学者对多个耦合网络的鲁棒性进行了分析，如通信与电力耦合网络，发现了不同于单个网络的渗流现象^[30-32]。

以上介绍了近年来通信领域复杂网络的研究热点，通信网络复杂属性的发现，如小世界特性和无标度特性，能更真实地反映通信网络实际运行规律，使得基于复杂网络的业务流行为研究更具参考价值。

2.5 本章小结

本章重点介绍了复杂网络的基本概念、度量参数以及典型的网络模型，为后续章节的开展奠定了基础。首先，介绍了网络的基本表示方法，包括图表示和邻接矩阵表示；其次，对复杂网络的基本度量参数进行了概述，包括度分布、平均最短路径长度、点/边介数、模块度等；再次，重点总结了目前经典的网络模型，包括最近邻耦合网络、ER 随机网络、WS 小世界网络和 BA 无标度网络。最后，总结了通信网络领域的复杂网络研究现状，包括网络结构、形成机制、鲁棒性等。本章简明扼要地介绍了与本文相关的复杂网络知识，方便各位读者更流畅的阅读后续内容。

参考文献

- [1] Barabási, A. L. Scale-free networks: a decade and beyond[J]. Science, 2009, 325(5939):412
- [2] Firth, J. A., Sheldon, B. C., Brent, L. Indirectly connected: simple social differences can explain the causes and apparent consequences of complex social network positions.[J]. Proc Biol Sci, 2017, 284(1867):20171939.
- [3] Li, Y., Wu, X., Li, L. Community Influence Analysis Based on Social Network Structures[A].// IEEE International Conference on Smart City/socialcom /sustaincom[C]. Chengdu: IEEE, 2016:247-254.
- [4] Safaei, A., Tavirani, M. R., Oskouei, A. A., et al. Protein-protein interaction network analysis of cirrhosis liver disease[J]. Gastroenterology & Hepatology from Bed to Bench, 2016, 9(2):114-123.

- [5] Zeng, H., Edwards, M. D., Liu, G., et al. Convolutional neural network architectures for predicting DNA-protein binding[J]. *Bioinformatics*, 2016, 32(12):121-127.
- [6] Barrat, A., Boccaletti, S., Caldarelli, G., et al. PREFACE:, Complex Networks: from Biology to Information Technology[J]. *Journal of Physics A Mathematical & Theoretical*, 2008, 41(22):220301.
- [7] Zhang, G. Q., Yang, Q. F., et al. Evolution of the Internet and its cores[J]. *New Journal of Physics*, 2008, 10(12):123027.
- [8] Qing, G. U., Xiong, S. J., Chen, D. X. Correlations between characteristics of maximum influence and degree distributions in software networks[J]. *Science China (Information Sciences)*, 2014, 57(7):1-12.
- [9] Lei, T. Research on computer network topology based on complex network theory[J]. *Computer Engineering & Applications*, 2007, 43(6):132-136.
- [10] Acharya, H. B., Chakravarty, S., Gosain, D. Few Throats to Choke: On the Current Structure of the Internet[A].// *Local Computer Networks*[C]. Singapore: IEEE, 2017:339-346.
- [11] Jin, S., Bestavros, A. Small-World Characteristics of the Internet and Multicast Scaling[A].// *Ieee/acm International Symposium on Modeling, Analysis and Simulation of Computer Telecommunications Systems*[C]. Orlando: IEEE, 2003:260-268.
- [12] Freeman, L. C. A set of measures of centrality based on betweenness[J]. *Sociometry*, 1977, 40(1):35-41.
- [13] Newman, M. E. J., Girvan M. Finding and evaluating community structure in networks[J]. *Phys Rev E Stat Nonlin Soft Matter Phys*, 2004, 69(2):026113.
- [14] Erdős, P., Rényi, A. On random graphs[J]. *Publicationes Mathematicae*, 1959, 6(4):290-297.
- [15] Watts, D., Strogatz, S. Collectivedynamics of 'small-world' networks[J]. *Nature*. 1998:440-442.

- [16] Newman, M. E. J., Watts, D. J. Renormalization group analysis of the small-world network model[J]. *Physics Letters A*, 1999, 263(4–6):341-346.
- [17] Barabási, A. L., Albert, R., Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509-512
- [18] Faloutsos, M., Faloutsos, P., Faloutsos, C. On power-law relationships of the Internet topology[J]. *Acm Sigcomm Computer Communication Review*, 1999, 29(4):251-262.
- [19] Yook, S. H., Jeong, H., Barabási, A. L. Modeling the Internet's large-scale topology [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99(21):13382-6.
- [20] Zhou, S., Mondragón, R. J. The rich-club phenomenon in the Internet topology[J]. *IEEE Communications Letters*, 2004, 8(3):180-182.
- [21] Bu, T., Towsley, D. On distinguishing between Internet power law topology generators[A].// in *Proceedings of Joint Conference of the IEEE Computer and Communications Societies [C]*. New York: IEEE, 2002:638-647.
- [22] Zhou, S., Mondragón, R. J. Towards modelling the internet topology—The interactive growth model [J]. *Teletraffic Science & Engineering*, 2003, 5:121-129.
- [23] Chang, H., Jamin, S., Willinger, W. Internet connectivity at the AS-level: an optimization-driven modeling approach[A].// *ACM SIGCOMM Workshop on Models, Methods and TOOLS for Reproducible Network Research[C]*. Karlsruhe: ACM, 2003:33-46.
- [24] Toyonaga, S., Kominami, D., Murata, M. Percolation analysis for constructing a robust modular topology based on a Binary-Dynamics model[J]. *International Journal of Distributed Sensor Networks*, 2017, 13(4):155014771770114.
- [25] Albert, R., Jeong, H., Barabási, A. L. Error and attack tolerance of complex networks[J]. *Nature*, 2000, 406(6794):378-382.
- [26] Wu, B., Tang, A., Wu, J. Modeling cascading failures in interdependent infrastructures under terrorist attacks[J]. *Reliability Engineering & System*

Safety, 2016, 147:1-8.

- [27] Liu, L., Yin, Y., Zhang, Z. Cascading failure of interdependent networks with traffic: Using a redundancy design to protect influential nodes[A].// International Conference on Reliability, Maintainability and Safety[C]. USA: IEEE, 2016:1-6.
- [28] Su, Z., Li, L., Peng, H., et al. Robustness of interrelated traffic networks to cascading failures[J]. Sci. Rep., 2014, 4:05413.
- [29] Dey, P., Mehra, R., Kazi, F., et al. Impact of Topology on the Propagation of Cascading Failure in Power Grid[J]. IEEE Transactions on Smart Grid, 2016, 7(4):1970-1978.
- [30] Buldyrev, S. V., Parshani, R., Paul, G., et al. Catastrophic cascade of failures in interdependent networks[J]. Nature, 2010, 464(7291):1025-1028.
- [31] Hong, S., Zhu, J., Braunstein, L. A., et al. Cascading failure and recovery of spatially interdependent networks[J]. Journal of Statistical Mechanics Theory & Experiment, 2017, 2017(10):103208.
- [32] Dai, Y., Chen, S., Cheng, Y. The study for protection strategy of cascading failure of interdependent network with the load[A].// Chinese Association of Automation[C]. Wuhan: IEEE, 2017:48-55.
- [33] 维基百科. 渗流模型 [EB]. <http://wiki.swarma.net/index.php/渗流模型>. 2016-08-21.

第三章 基于复杂网络的业务流服务质量 耦合行为建模与机理研究

全球网络承载了数以亿计的通信业务，并为每项业务提供满足其传输需求的基础底层设施。具体而言，每条业务流从源节点出发将经过多个网络单元（路由器、链路）到达目的节点，每个网络单元又同时承载多条业务流，这使得全网业务流之间呈现明显的关联关系^[1-2]，本文将此相关性称之为耦合性。由于分布式的网络资源十分有限，业务间的耦合关联尤其是服务质量之间的竞争关系变得更为明显^[3-6]，本章将针对业务流服务质量间的关联性进行数学建模，并挖掘业务服务质量间的耦合机理，提升业务传输质量。

一些学者针对网络中存在的业务关联现象开展了大量研究，大多集中于业务流到达的相关性，如短相关特性^[7-8]、长相关特性^[9-10]以及业务的自相似性^[11-13]等，这些研究结果并不直接关联于业务的服务质量，如时延、丢包率等，因此尚不能有效应用于网络业务服务质量的提升。除此之外，Barthelemy 等人^[14]使用随机矩阵理论（Random Matrix Theory, RMT）分析了物理链路间的关联关系，相关结果显示链路间的强相关可能是造成流量自相似特性的原因之一，网络中路由器之间不同的连接方式对业务集体行为的表现起着至关重要的作用。文献^[14]同时揭示了网络中“active center”的存在，“active center”与大量路由器交换信息，从而引发相应连接之间的相关性。Wu 等人^[15]基于 Internet 实际流量数据（P2P、HTTP、即时消息等）与业务流/服务器连接偏好建立了业务流连接关系图（Flow Graph），并仿真得到 flow graph 满足幂律分布，且不同业务类型具有显著不同的幂律指数。此外，文献^[15]考察了业务的增长特性，业务的动态增长属性取决于现有业务的分布特点，并不由新到达的业务支配。可以看出，目前业务流关联性成果主要包括了时间尺度上的短/长相关性、业务流自相似性、物理链路的强相关性以及业务自身增长属性等，尚未涉及到服务质量相关性方面，关于业务流间服务质量耦合关联的研究仍为空白。

本章针对大规模复杂网络中业务流服务质量的关联性进行建模分析，拟从微观耦合机理与宏观关联关系两方面共同刻画全网业务服务质量之间的关联属性。

微观上，建立业务流弹性物理模型，从物理角度刻画业务服务质量之间的弹性耦合机理。具体而言，利用弹性力学相关理论知识，构建通信网络系统与弹性物理系统的映射关系，将业务流服务质量间的关联耦合作用类比于弹性体间的相互作用，建立业务流弹性物理模型。利用该模型，分析业务流服务质量耦合关联的内在原因，并结合弹性力学中势能的概念，提出一种度量网络性能优劣的新指标。

宏观上，基于业务流的微观弹性耦合作用，构建全网业务之间的耦合关联网络，研究大规模业务流在通信网络中的分布特点、聚集特点与耦合关联特点。具体而言，利用复杂网络度量分析手段，从特征参数度分布、平均路径长度、聚集系数入手，挖掘业务流耦合关联网络结构特征并分析特征参数的实际通信含义。

3.1 业务流弹性物理模型

本节引入弹性力学知识，构建通信系统和物理系统的映射关系，借助弹性理论建立业务流服务质量耦合关联模型。

3.1.1 弹性理论引入

首先对弹性力学中几个重要的物理量以及弹性系统的构成进行介绍。弹性系统由弹性体和弹性体间的相互作用构成^[16-17]。弹性体在受到外力作用时会产生形变，当恢复至初始状态时，将对系统中其他物体产生力的作用，弹性体之间通过力的作用可使系统维持稳态。

(1) 应力 σ

物体由于外力作用产生形变，在物体各部分之间产生相互作用的内力称作应力。应力反映了物体力图恢复到初始位置状态的趋势，公式表示为：

$$\sigma = \lim_{\Delta S_c \rightarrow 0} \frac{\Delta p}{\Delta S_c} \quad (3-1)$$

Δp 是作用在微小面积 ΔS_c 上的内力矢量。

如图 3-1 所示，应力可以分解为两个分量：所在平面的外法线方向和切线方向。外法线方向的应力分量叫做正应力，记作 σ_n 。沿切线方向的应力分量叫做剪应力，记作 τ_n 。

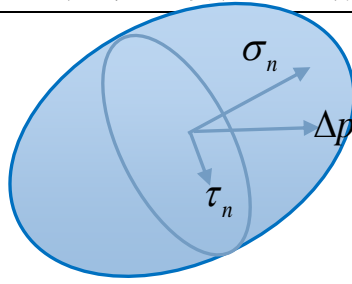


图 3-1 应力

(2) 应变 ε

物体内任意点因外力作用引起的形状和尺寸的相对改变称作应变。应变分为线应变、剪应变和体应变。

线应变：物体上某微小线段在形变后长度的相对变化 $\frac{dl}{l}$ ；

剪应变：物体上相互垂直的微小线段形变后所夹角度的变化 $\frac{d\phi}{\phi}$ ；

体应变：物体形变后体内任一微小单元体积 V 的改变，表示为：

$$\frac{dV}{V} = \frac{2\pi r l}{\pi r^2 l} \frac{dr}{r} = \frac{2}{r} \frac{dr}{r} \quad (3-2)$$

(3) 应力应变关系

应力和应变的数学关系可以通过广义胡克定律进行表示^[17]：

$$\sigma = E\varepsilon \quad (3-3)$$

其中， E 为弹性模量。弹性模量分为杨氏模量 E 和剪切模量 G 。杨氏模量为正应力与线应变之比值；剪切模量为剪应力与剪应变之比值。对于各向同性、均匀的弹性体，两者的关系为：

$$G = \frac{E}{2(1+\mu)} \quad (3-4)$$

其中， μ 为泊松系数。对同一种材料，在弹性极限内，弹性模量为常数。

(4) 应变能

弹性体受外力作用产生形变后，在弹性体内存储的势能称作应变能。外力缓慢地加到物体上时可视作静力，由于物体不会因静力作用产生运动，因此可忽略

动能以及其他能量的消耗，外力势能全部转化为弹性体的势能。据此应变能的表达式为：

$$U = \iiint_V U_0 \, dV \quad (3-5)$$

其中， U_0 为单位体积的应变能，表示为：

$$U_0 = \int_0^{\varepsilon_x} \sigma_x d\varepsilon_x = \frac{1}{2} E \varepsilon_x^2 = \frac{1}{2} \sigma_x \varepsilon_x \quad (3-6)$$

(5) 弹性势能

弹性势能是在物体发生弹性形变后，物体内部各部分由于内力相互作用产生的势能。如果规定弹性体未发生形变时的势能为“势能零点”，那么外力所做的功等于物体形变达到 Δl 时的势能：

$$\begin{aligned} P &= \int_0^{\Delta l} ES \frac{\varepsilon}{l_0} d\varepsilon = \frac{1}{2} \frac{ES}{l_0} \varepsilon^2 \Big|_0^{\Delta l} \\ &= \frac{1}{2} E \left(\frac{\Delta l}{l_0} \right)^2 l_0 S = \frac{1}{2} E \left(\frac{\Delta l}{l_0} \right)^2 V \end{aligned} \quad (3-7)$$

依据弹性体应变的定义，上式可化简为：

$$P = \frac{1}{2} E \varepsilon^2 V \quad (3-8)$$

若弹性体为均匀弹性体，那么弹性势能密度为：

$$P^0 = \frac{1}{2} E \varepsilon^2 \quad (3-9)$$

以上对弹性力学的基础概念进行了简单介绍，涉及的物理参数将在后文中与通信网络参数一一对应，并通过公式证明完成弹性系统与通信系统的映射。

3.1.2 模型建立

首先，简述业务流服务质量间的弹性作用机制。当网络中某条业务流的流量突然增加时，必将对与其共享链路的其他业务流造成影响，使得邻居流的带宽遭受挤压；而当某条业务流流量减少时，又使得邻居流被挤压占用的带宽得到恢复。本文把业务流带宽随网络环境不断变化的特性称为“业务流弹性”，并将业务流服务质量间的耦合关联性与物理系统中弹性体的相互作用进行类比。

从单位时间来看，如果将任一业务流类比于弹性体，那么弹性体的横截面积就正比于业务流的流量大小，即带宽。当业务流单位时间内流量增加时（带宽增加），对应于弹性体的横截面积相应增加，此时弹性体产生应变，向外产生应力，释放应变能；其邻居流由于受到挤压作用，产生相应的应力和应变，带宽减少，储存应变能。

以上过程需满足如下条件：如果认为网络实际提供的链路带宽总量 C 为一个刚性管道（不可变），用户“需求” D 是一个固定体积的弹性体，要把固定体积的弹性体放进管道中就有可能因为得不到满足而产生向外扩张的压力，即每一个弹性体发生应变是由于供需关系不平衡导致的（ $C \leq \sum D$ ）。此时，每条业务流的受力结果为自身产生应力与所有邻居流产生应力的合力，而仅有当两种应力相抵消，即合力为 0 时，业务流的分布才处于稳定状态。

表 3-1 列出了通信系统与弹性物理系统的参数映射关系。通过分析业务流服务质量相互作用机制与弹性物理系统的关系，建立了业务流的弹性物理耦合模型。

模型假设所有链路容量及允许的流量上限都一致；同一路径上，双向业务流具有对称性，所有属性相同；固定路由集，暂不考虑次优路径；每条业务流的实际流量仅受到其带宽需求和链路容量的限制；作为扩展，可以继续考虑其它 QoS 参数的制约，如时延、丢包率。

K 表示业务流 f_i 经过的链路集。业务流 f_i 在其路径的第 k 条链路（ $k \in K$ ）中的弹性特性描述如下：

$$\sigma_i^k = \varphi_i \times \varepsilon_i^k \quad (3-10)$$

公式（3-10）基于广义胡克定律^[17] $\sigma = E \times \varepsilon$ 。应力 σ 表示弹性体间的作用力大小， E 表示弹性体的弹性模量，应变 ε 表示弹性体由于外力作用导致的几何尺寸相对变化量。在这里，将业务流的通信特性与弹性力学特性进行对应，应力 σ_i^k 表征业务流 f_i 对第 k 条链路中其他业务流的作用大小； φ_i 表示业务流的传输优先级，优先级高的业务带宽优先被保证，不易受到外界影响，优先级低业务带宽易被挤压，受外界环境影响大。业务流应变 ε_i^k 表征业务在第 k 条链路中带宽受挤压的程度，可通过下式计算得到：

$$\varepsilon_i^k = \frac{D_i - B_i^k}{D_i} \quad (3-11)$$

表 3-1 物理系统与通信网络系统类比表

参数	通信系统	符号	物理系统	符号	注释
业务流	不同 QoS 等级的业务流	F	线性弹性体	N	///
业务流特性	用户需求	D	想压入管道的单位长度体积	V_d	只有当 $V_d \geq V_b$ 时, 才会发生应变
	实际流量	B	单位长度的实际允许体积	V_b	
链路容量	链路的承载能力	C	管道总横截面积	S	管道为刚体, 总体积不变
应变	业务流带宽增加或减小的相对变化	$\frac{D-B}{D}$	弹性体横截面积的变化	ε	当产生挤压时, 弹性体发生应变
应力	业务流恢复其带宽的趋势大小	///	弹性体抵抗应变的力	σ	///
弹性模量	业务优先级	φ	形变的难易程度	E	在弹性形变范围内为常数
弹性势能	业务流在受挤压过程中的拥塞状态	P	弹性体形变后储存在体内的能量	U	///

其中, D_i 为业务需求带宽, B_i^k 为业务实际传输带宽 ($B_i^k \leq D_i$), $D_i - B_i^k$ 表示单位时间内未被传输的业务量, 即拥塞业务量。

在链路拥塞状态下, 链路 k 中存在以下关系:

$$\begin{cases} \sum B_i^k = C_k \\ \sum D_i \geq C_k \end{cases} \quad (3-12)$$

其中, C_k 是链路 k 的容量。

在链路空闲状态下, 链路 k 中存在以下关系:

$$\begin{cases} \sum B_i^k < C_k \\ \sum D_i < C_k \end{cases} \quad (3-13)$$

依据弹性力学知识, 网络存在业务流耦合作用的稳态, 即业务流间的相互作用力平衡时, 业务带宽分布达到稳态:

$$\sigma_1^k = \sigma_2^k = \dots = \sigma_i^k = \dots = \sigma^k \quad (3-14)$$

稳态证明详见 3.1.3 节“模型证明”。值得注意的是, 在链路空闲状态下, 应力满足下式:

$$\sigma_1^k = \sigma_2^k = \dots = \sigma_i^k = \dots = 0 \quad (3-15)$$

联合公式(3-10)至(3-15), 可以得到业务 f_i 在第 k 条链路中的传输带宽 B_i^k 。在实际网络中, 不同的应力平衡态导致链路的拥塞状态也不同, 业务可允许传输的最大带宽也将不同。考虑到业务流带宽在传输路径上的一致性, 业务最终带宽取决于最拥挤的链路, 即应力和应变最大的链路, 因此, 业务 f_i 的最终带宽为:

$$B_i = \min(B_i^1, B_i^2, \dots, B_i^k, \dots) \quad (3-16)$$

利用相同的方法, 可以得到全网所有业务流传输平衡状态下的带宽。

此外, 对业务流在传输过程中的能量状态进行分析。弹性势能反映了弹性体因受到挤压作用体内存储能量情况。在弹性形变范围内, 弹性体形变越大, 具有的弹性势能就越多; 反之, 弹性势能则越小。对应到通信系统, 业务流带宽受挤压实际反映了业务流的拥塞情况, 可视作网络能量的蓄积过程。因此, 基于弹性势能概念定义网络“弹性拥塞系数”。依据公式 (3-8), 业务流 f_i 的弹性拥塞系

数为:

$$P_i = \frac{1}{2} D \varphi \varepsilon_{i, \max}^2 \quad (3-17)$$

其中, 应力 $\varepsilon_{i, \max}$ 满足:

$$\varepsilon_{i, \max} = \frac{D_i - B_i}{D_i} \quad (3-18)$$

综上, 通过建立业务服务质量间的弹性物理模型 (以带宽为例), 依据业务间弹性耦合平衡状态, 可以得到网络中任一业务的带宽分布和拥塞状态。

3.1.3 模型证明

(1) 情况一: 静态饱和链路中应力稳态证明

给定饱和链路容量为 C , 链路中业务流 f_i 的带宽需求为 $D_i(t)$, 实际传输带宽为 $B_i(t)$ ($0 \leq B_i(t) \leq D_i(t)$), 业务流 f_i 的带宽变化来自两方面的作用: 一方面是自身受到挤压产生的反作用力; 另一方面是邻居流对业务流 f_i 的合力。两种作用力方向相反, 当两者平衡时, 业务流 f_i 带宽达到稳定状态。

依据胡克定律^[17], 定义业务流 f_i 的流量变化率:

$$\frac{d \ln B_i(t)}{dt} = E_i \cdot \frac{D_i(t) - B_i(t)}{B_i(t)} - E_j \cdot \frac{\sum_{j \neq i} D_j(t) - \sum_{j \neq i} B_j(t)}{\sum_{j \neq i} B_j(t)} \quad (3-19)$$

其中, E_i 为业务流 f_i 的弹性模量, E_j 为邻居流的平均弹性模量。在无差别的网络中, 业务流的弹性模量 (即业务优先级) 均相同, $E_i = E_j = 1$, 等式 (3-19) 可化简为:

$$\frac{dB_i(t)}{dt} = D_i(t) - B_i(t) - \frac{\sum_{j \neq i} D_j(t) - \sum_{j \neq i} B_j(t)}{\sum_{j \neq i} B_j(t)} \cdot B_i(t) \quad (3-20)$$

公式 (3-20) 两侧同时对 i 求和可得:

$$\sum_i \frac{dB_i(t)}{dt} = \sum_i D_i(t) - \sum_i B_i(t) - \sum_i \frac{\sum_{j \neq i} D_j(t) - \sum_{j \neq i} B_j(t)}{\sum_{j \neq i} B_j(t)} \cdot B_i(t) \quad (3-21)$$

表 3-2 符号说明

物理量	含义
f_i	链路中的业务流 ($i=1,2,3,\dots$)
$D_i(t)$	t 时刻业务带宽需求
$B_i(t)$	t 时刻业务实际带宽
C	链路容量
E_i	业务流 f_i 的弹性模量
E_j	链路中其它流的平均弹性模量
φ_i	业务 f_i 优先级
k	链路单位面积承载的实际流量, 为常数
U	网络业务的应变能
μ	单位流量对应的应变能
$U_0(t)$	业务流的应变能密度
$S(t)$	业务流在链路中对应的面积

公式 (3-21) 左侧表示链路中业务流的带宽变化之和。在链路饱和状态下 ($\sum_i B_i(t) = C$), 带宽变化之和为零, 故公式 (3-21) 化简为:

$$\begin{aligned}
 0 &= \sum_i D_i(t) - C - \sum_i \frac{\sum_{j \neq i} D_j(t) - \sum_{j \neq i} B_j(t)}{\sum_{j \neq i} B_j(t)} \cdot B_i(t) \\
 &= \sum_i D_i(t) - \sum_i \frac{\sum_{j \neq i} D_j(t)}{\sum_{j \neq i} B_j(t)} \cdot B_i(t)
 \end{aligned} \tag{3-22}$$

在稳定的无差别网络中 (业务优先级相等), 存在业务带宽需求 $D_i(t)$ 与实际传输带宽 $B_i(t)$ 的关系式:

$$\frac{D_1(t)}{B_1(t)} = \dots = \frac{D_i(t)}{B_i(t)} = \dots = \frac{D_n(t)}{B_n(t)} = \text{常数} \quad (3-23)$$

依据公式 (3-23)，公式 (3-22) 进一步推导为：

$$\begin{aligned} 0 &= \sum_i D_i(t) - \sum_i \frac{\sum_{j \neq i} D_j(t)}{\sum_{j \neq i} B_j(t)} \cdot B_i(t) \\ &= \sum_i D_i(t) - \sum_i \frac{D_i(t)}{B_i(t)} \cdot B_i(t) \\ &= 0 \end{aligned} \quad (3-24)$$

最终，推导证明公式 (3-21) 等号成立。公式 (3-21) 左侧表示饱和链路中所有业务的带宽变化之和，右侧表示链路平衡状态下业务流之间的应力之和。公式 (3-21) 表明，当链路中应力相互抵消，即应力稳态求和为零时，链路中所有业务的带宽变化之和也为零，带宽分布达到平衡。

对于更普遍的存在业务流优先级的有差别网络，业务带宽需求 $D_i(t)$ 与实际传输带宽 $B_i(t)$ 有如下关系：

$$\varphi_1 \frac{D_1(t)}{B_1(t)} = \dots = \varphi_i \frac{D_i(t)}{B_i(t)} = \dots = \varphi_n \frac{D_n(t)}{B_n(t)} = \text{常数} \quad (3-25)$$

此时，由于优先级系数的存在，网络中业务流对带宽的需求程度将不同，即业务基本需求 $D_i(t)$ 不能够完全反映业务实际的带宽需求。因此，重新定义业务实际需求：

$$D'_i(t) = \varphi_i D_i(t) \quad (3-26)$$

利用业务实际需求 $D'_i(t)$ 代替基本需求 $D_i(t)$ ，从而将有差别网络转化为对应的无差别网络，则应力稳态证明过程与无差别网络相同。

(2) 情况二：动态环境下应力稳态证明

当链路中某条业务流带宽发生改变时，将影响其所在链路中其他的业务流，从而引起整个网络中业务带宽的重新分配。利用数学公式刻画这一过程，首先定义应变能密度：

$$U_0(t) = \frac{U(D) - U(B)}{S(t)} \quad (3-27)$$

其中, U 表示业务流 f_i 的应变能, $S(t)$ 表示业务流 f_i 的横截面积大小。令 μ 为单位流量对应的应变能, k 为单位横截面积承载的实际流量, 三者存在以下关系:

$$\begin{cases} U(D) = \mu \cdot D(t) \\ U(B) = \mu \cdot B(t) \\ B(t) = k \cdot S(t) \end{cases} \quad (3-28)$$

依据公式 (3-28), 应变能密度 $U_0(t)$ 可化简为:

$$U_0(t) = k\mu \left(\frac{D(t)}{B(t)} - 1 \right) \quad (3-29)$$

依据应变能与应力、应变的关系, $U_0(t)$ 同时也满足下式:

$$U_0(t) = \frac{1}{2} \sigma(t) \varepsilon(t) \quad (3-30)$$

依据公式 (3-30) 应变能密度随时间变化率为:

$$\frac{dU_0(t)}{dt} = \frac{1}{2} (\sigma(t) \varepsilon(t))' = \frac{\sigma(t)}{\varphi} \quad (3-31)$$

依据公式 (3-29) 应变能密度变化率也可表示为:

$$\begin{aligned} \frac{dU_0(t)}{dt} &= k\mu \left(\frac{D(t)}{B(t)} - 1 \right)' \\ &= k\mu \frac{D'(t)B(t) - B'(t)D(t)}{B^2(t)} \end{aligned} \quad (3-32)$$

由于业务需求 $D(t)$ 通常为常数, 公式 (3-32) 化简为:

$$\begin{aligned} \frac{dU_0(t)}{dt} &= -k\mu D(t) \frac{B'(t)}{B^2(t)} \\ &= -k\mu D(t) \frac{kS'(t)}{k^2 S^2(t)} \\ &= \mu D(t) \left(\frac{1}{S(t)} \right)' \end{aligned} \quad (3-33)$$

在无差别网络中，业务弹性模量 φ 相等，对比公式 (3-31) 与公式 (3-33)，

令 $\varphi = \frac{1}{\mu D(t)}$ ，那么应力 $\sigma(t)$ 为：

$$\sigma(t) = \left(\frac{1}{S(t)} \right)' \quad (3-34)$$

在无差别网络中，任一业务流对应到链路中所占的面积 $S(t)$ 均相等，因此，应力 $\sigma(t)$ 为一常数，即动态网络环境下，当应力相等时，链路业务的带宽分布处于平衡状态。

在有差别网络中，依据公式 (3-26) 重新定义业务实际需求 $D_i'(t) = \varphi_i D_i(t)$ 。优先级系数虽影响弹性模量的变化，但并不影响应力的平衡状态，依然满足应力相等时，链路中业务流带宽分布达到稳定状态。

3.2 业务流耦合关联网络

将业务间弹性耦合作用推广至整个网络，建立无向带权重的业务流耦合关联网络 $G(V, E)$ 。该网络中，业务视作节点，节点集合 V 代表网络中的业务流；业务间弹性耦合作用视作边，边集合 $E \in V \times V$ 表示业务流间的弹性耦合关系，存在共享链路的两个业务之间建立连边。 $W(G) = (w_{i,j})_{N \times N}$ 为边权重矩阵。若业务 i 与业务 j 之间存在连边，则赋值 $w_{i,j} = \sigma$ ；否则， $w_{i,j} = 0$ 。 σ 是业务间弹性作用力的大小，由上节建立的业务流弹性物理模型计算得出。

业务流耦合关联网络（Flow Interaction Network）的具体构建步骤如下：

Step1: 构建通信网络拓扑，如图 3-2 (a) 所示。由于网络层可以直接反映业务的传输情况，如传输路径、带宽需求、服务等级等，因此本文使用路由级拓扑；

Step2: 获取业务流路由集合 R 。仅考虑固定路由集，业务均选择最短路径，暂不考虑次优路径；

Step 3: 业务关联关系建立。

考察任意两条业务的路由集 R_i 和 R_j ，若 $R_i \cap R_j \neq \emptyset$ ，那么节点 i 与节点 j 之间建立连接；否则，节点 i 与节点 j 之间无连接。举例说明，业务 1 与

业务 2 的路径交集不为空集（见图 3-2（a）），因此图 3-2（b）中节点 1 与节点 2 之间存在一条连接。

Step4: 重复 Step 3，直至考察完所有业务对，完成业务流耦合关联网络的构建。

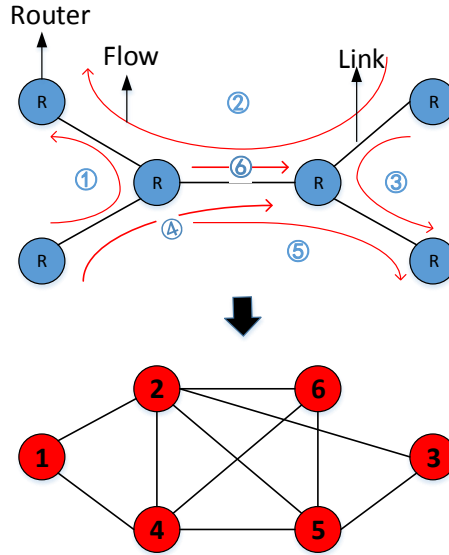


图 3-2 （a）网络及业务流分布；（b）业务流耦合关联网络。（a）中节点表示路由器，（b）中节点表示业务流，（b）中的节点编号与（a）中业务流编号一致。

业务流耦合关联网络是业务间弹性耦合作用的图形化表示，可直观展示网络中所有业务流之间的关联关系，宏观上挖掘业务关联性的整体特征，并分析业务间相互影响的传播特点。

3.3 实验仿真与结果分析

3.3.1 网络环境与参数设置

网络设置：20 个路由节点构成的随机网络，链路容量为 $C=5$ (Gbps)，并保证任意两个节点之间都存在路径。

业务设置：业务流数目由 0 到 500 逐步增加，每条业务流随机选择起始节点和目的节点，按照 Dijaska 算法进行选路。业务的带宽需求满足正态分布 $D \sim N(2, 0.3)$ 。在无差别网络中，业务的优先级设定为 $\varphi=1$ ；在有差别网络中，业务的优先级分为 3 个等级 $\varphi=1, 2, 3$ ，并随机分配业务优先级。

3.3.2 业务流小世界特性

依据上节网络参数设置,可构建得到业务流从 0 增长至 500 的动态耦合关联网络,完成网络拓扑到业务逻辑拓扑的映射。基于该动态拓扑,从复杂网络角度仿真业务流总体分布特点,包括业务度分布、业务间平均路径长度、业务聚集系数等。通过分析网络结构参数特征,挖掘业务服务质量间的宏观关联属性。

业务流耦合关联网络为加权网络,加权网络的边权值将影响节点间的聚集程度,因此,将传统聚集系数计算公式(2-8)进行推演,得到业务流耦合关联网络的聚集系数公式:

$$C_i = \frac{w_i n}{C_{k_i}^2} = \frac{2w_i n}{k_i(k_i - 1)} \quad (3-35)$$

其中, w_i 为归一化加权系数,表示为:

$$w_i = \frac{\bar{w}}{w_{\max}} \quad (3-36)$$

其中, \bar{w} 为业务 i 与邻居流之间的平均边权值, w_{\max} 为最大边权值。

设置具有相同节点数的随机网络作为比较。图 3-3 分别显示了业务流数目从 0 增长到 500 过程中随机网络和业务流耦合关联网络的平均路径长度和聚集系数的变化情况。图 3-3 (a) 为网络平均路径长度的变化。当业务流数目小于 40 时,业务流耦合关联网络和随机网络的平均最短路径长度差别较小;随着业务流数目的增长,业务流耦合关联网络的平均最短路径长度逐渐稍大于随机网络,符合小世界网络判定的第一特征:小世界网络的平均路径长度稍大于随机网络。图 3-3 (b) 显示了聚集系数的变化。可以看出,随着业务流数目的增多,业务流耦合关联网络的聚集系数逐渐远大于随机网络,且最终趋于稳定。前者聚集系数达到了 0.6,而后者只有 0.1 左右,符合小世界网络判定的第二个特征:小世界网络的聚集系数远大于随机网络。综合图 3-3 (a) 和 3-3 (b) 两项结果,可以判断业务流耦合关联网络具有小世界特征:平均最短路径长度较短(本文实验环境中小于 5);聚集系数较大(本文实验环境中达 0.6)。

通过以上分析得出结论:当网络中的业务流随机分布且数目增加到一定程度时,业务流耦合关联网络呈现小世界特征:1) 一条业务可以通过有限次的流间耦合关系影响到另一条业务;2) 随着业务数量的增长,网络业务间的关联程度

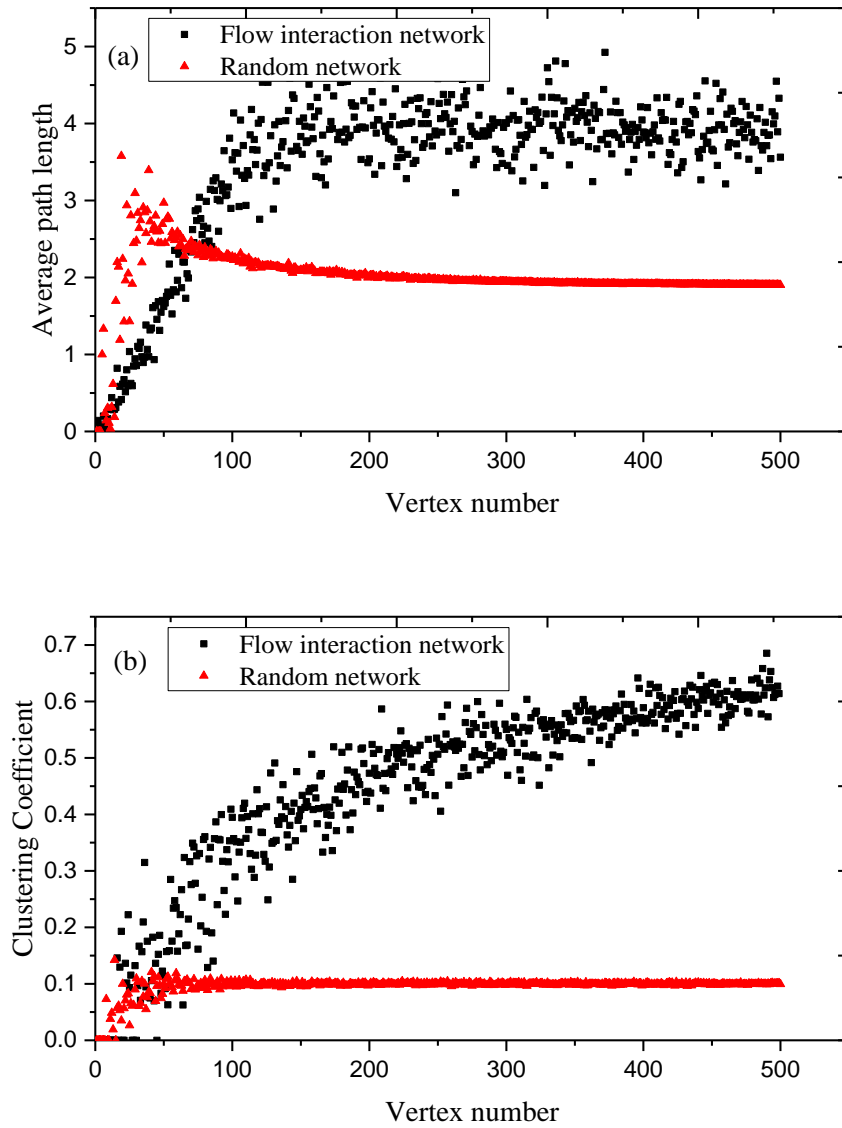


图 3-3 (a) 平均路径长度; (b) 平均聚集系数

最终趋于恒定，并呈现明显的聚集特征，表明业务间密切的弹性竞争关系。

此外，对业务流耦合关联网络中的业务度分布进行了统计分析，并与具有相同节点数和平均度的随机网络进行了比较。已有研究表明，随机网络度分布服从二项分布：

$$P(k) = C_{N-1}^k p^k (1-p)^{N-1-k} \quad (3-37)$$

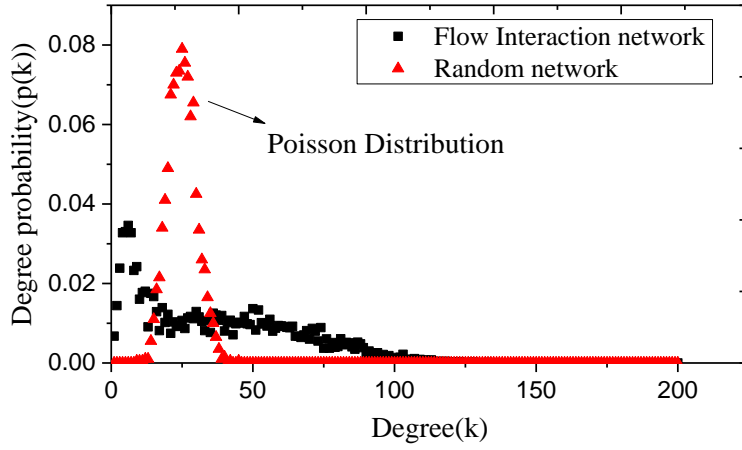


图 3-4 业务流关联耦合网络度分布和随机网络度分布

当 N 充分大时，度分布函数近似为泊松分布^[18]：

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (3-38)$$

其中， $\langle k \rangle = pN$ ， p 为节点间随机连接概率。

小世界网络的度分布理论上类似于泊松分布^[19]：

$$P(k) = e^{-mp} \frac{(mp)^{k-m}}{(k-m)!} \quad (3-39)$$

其中 m 为小世界模型的初始节点度数， p 为节点之间添加“长程连接”的概率。

图 3-4 显示了具有 1000 条流的业务流耦合关联网络和 1000 个节点的随机网络度分布情况。可以看出，业务流耦合关联网络的度分布随着度数增加开始迅速增加然后缓慢减小。当 k 值较小时 ($0 < k < 20$)，度分布基本符合泊松分布，出现一个类似山峰状的凸起；当 k 较大时 ($20 < k < 120$)，则显现出类似长尾分布的现象，与小世界网络的理论度分布有一定偏差。事实上， k 值较大时出现的类似长尾分布的现象是由于网络资源利用率极不均衡造成的：异常拥塞的链路中业务量通常远远高于低利用率的链路，而异常拥塞的链路通常只占网络全部链路的小部分，造成了异常拥堵的链路中业务流的度数极高，总体数目却很少的现象。

以上解释了业务流耦合关联网络中出现了 $100 < k < 120$ 的高连接度业务但概率值却极小接近于 0 的原因。业务流的度数可表征影响范围的大小，图 3-4 的结果表明，网络中小部分的业务流有极为庞大的影响范围，称之为网络中的“超级流”。

图 3-4 显示了具有相同节点数的随机网络的度分布情况。可以看出，随机网络的度分布呈现出离散泊松分布样式，符合公式 (3-39) 的理论描述。图 3-4 中，随机网络大部分节点的度数集中在平均值 25 左右，最高度数仅为 45。随机网络的节点度分布差异性较小，而业务流耦合关联网络的度分布差异却非常大。通过对比实验，发现业务流耦合关联网络的度分布明显不同普通网络。

3.3.3 业务流拥塞行为评估

发生弹性形变的物体在体内储存弹性势能，同一弹性体形变越大，具有弹性势能越多。对应到通信系统，业务流受挤压蓄积能量的过程实际上反映了业务流的拥塞情况，因此，依据弹性势能定义了“弹性拥塞系数”（公式 (3-17)）。通过将网络中所有业务流的拥塞系数求和，可以得到整个网络的拥塞情况。以往研究表明，当业务流分别具有相同的优先级和比例优先级时，网络的拥塞情况不同。如果业务拥有相同的优先级，网络资源利用率和网络吞吐量优于比例优先级，即网络越公平，资源分配越优化^[20-22]。因此，业务具有相同优先级时网络的拥塞程度将小于比例优先级。基于上述研究结果，将“弹性拥塞系数”与网络丢包率指标进行对比实验，比较两者在评估网络拥塞时的区别和性能优劣。

实验环境：通信网络为 10~90 个节点的随机网络，网络动态变化，节点增加步长为 10。网络链路容量为 $C = 5$ (Gbps)，业务流总数为 300，每条业务流随机选择源节点和目的节点，按照 Dijaska 算法进行选路。业务的带宽需求满足正态分布 $D \sim N(3, 0.4)$ ，构建拥塞的网络环境。考虑两种网络业务：1) 业务流具有相同优先级 $\varphi=1$ ；2) 业务流具有比例优先级 $\varphi=1, 2, 3$ ，并随机分配业务优先级。

图 3-5 (a) 显示了利用平均丢包率评估网络拥塞的结果。可以观察到，在网络业务流数目和业务带宽需求不变的情况下，随着网络规模的增加，链路资源不断丰富，网络业务丢包率逐渐缓慢下降。图 3-5 (b) 显示了网络业务平均“弹性拥塞系数”在相同网络环境下的仿真结果。可以观察到，随着网络规模的增加，“弹性拥塞系数”逐渐下降，呈现出类似于图 3-5 (a) 的下降趋势（业务具有比例优先级时）。因此，对于一般的支持比例优先级业务的网络，两个指标可类比，即“弹性拥塞系数”可用于评估业务平均拥塞状态。

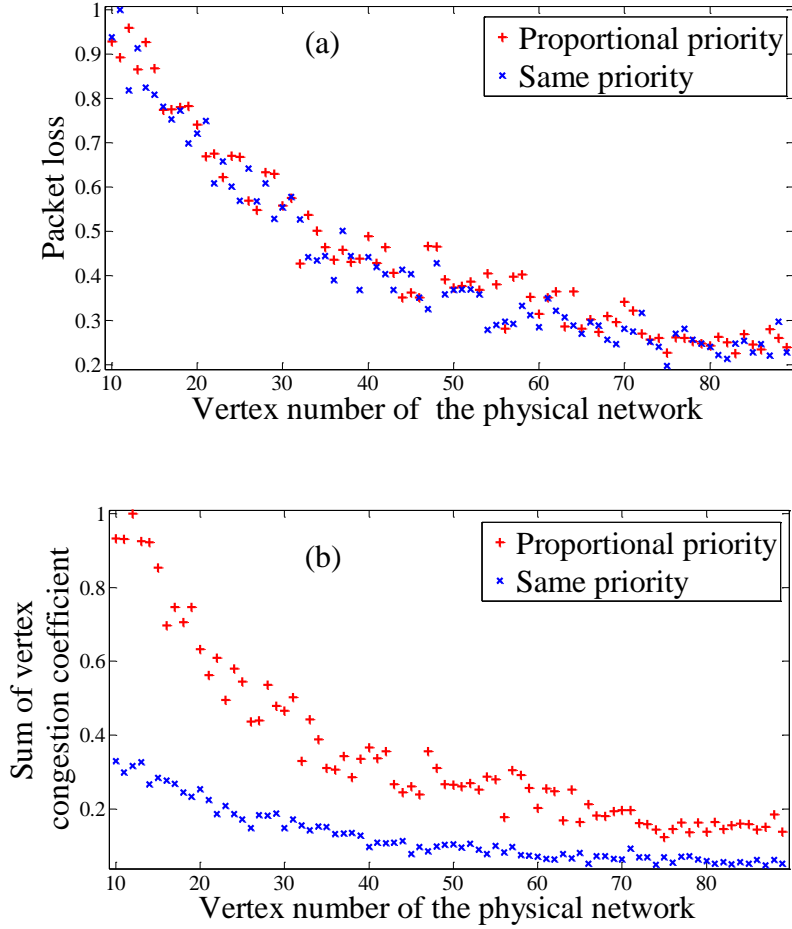


图 3-5 (a) 网络平均丢包率; (b) 弹性拥塞系数。

注：数值进行了归一化处理

另外，图 3-5 (a) 显示出业务具有相同优先级时网络平均丢包率较小、拥塞程度较低，但大量的数据混合难以明显分辨同优先级业务和比例优先级业务之间的拥塞差异。而图 3-5 (b) 中同优先级业务和比例优先级业务之间的“弹性拥塞系数”差异更为明显：当业务具有相同的优先级时，业务“弹性拥塞系数”更小，网络拥塞程度明显更低，能够反映网络实际传输情况。“弹性拥塞系数”考虑了业务优先级以及带宽传输需求的不同，使得不同优先级业务的拥塞差异更明显。由此可见，在本文实验环境下，“弹性拥塞系数”相比于传统的丢包率指标更能真实地反映网络拥塞情况。

业务流弹性拥塞系数的概率分布如图 3-6 所示，横坐标为业务弹性拥塞系数，纵坐标为概率值，数值均进行了归一化处理，并使用 \log 值表示。可以看出，业务弹性拥塞系数的概率分布大致满足幂律分布，即数据进行 \log 处理后满足线性

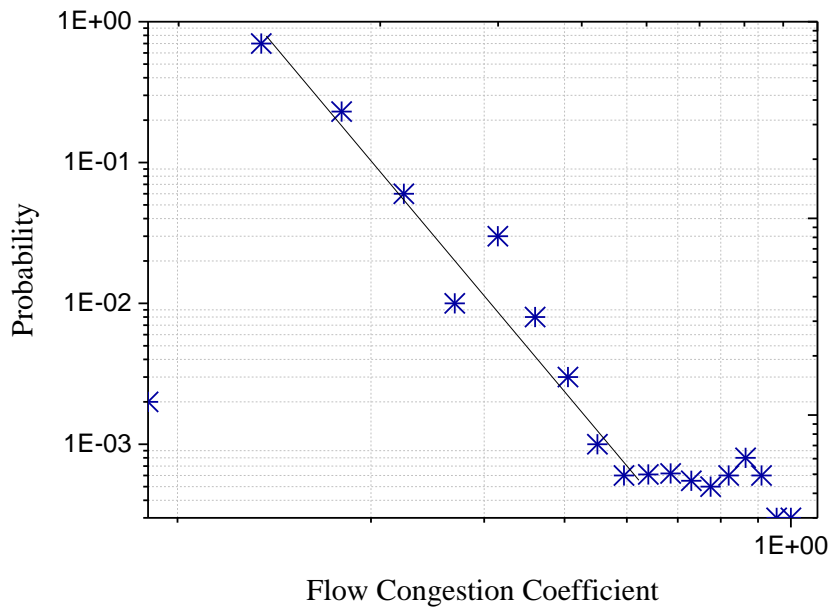


图 3-6 业务弹性拥塞系数概率分布

关系。依据幂律分布特性，可以得出以下结论：在网络拥塞环境下，大多数的业务处于轻度拥塞的状态，仅有少部分的业务处于极度拥塞的状态。

3.4 本章小结

本章建立了业务流服务质量间的弹性物理耦合模型，并基于业务间弹性耦合作用构建了无向加权的业务流耦合关联网络：节点表示业务流，连边表示业务流间的弹性关联。在业务流耦合关联网络中，首先对其结构特征进行了研究，实验结果发现网络中随机分布的业务流在达到一定数目时表现出了小世界特性：聚集系数高、平均最短路径长度小，有助于从宏观上理解业务间的耦合作用模式。其次，业务流耦合关联网络的度分布证明了网络中“超级流”的存在，超级流在网络中具有广泛的影响力。最后，基于业务流弹性物理模型提出了“弹性拥塞系数”的概念。“弹性拥塞系数”一方面可以反映网络拥塞情况，另一方面可以反映不同优先级业务的拥塞差异。业务流弹性拥塞系数的概率分布在网络拥塞环境下大致满足幂律分布，表明大部分业务的拥塞程度较低，仅有少数业务处于严重拥塞状态。

参考文献

- [1] Jin, H., Yao, X., Chen, Y. Correlation-aware QoS modeling and manufacturing cloud service composition[J]. Journal of Intelligent Manufacturing, 2017, 28(8):1947-1960.
- [2] Huang, J., Li, S., Duan, Q., et al. QoS Correlation-Aware Service Composition for Unified Network-Cloud Service Provisioning[A].// Global Communications Conference[C]. Washington: IEEE, 2017:1-6.
- [3] Gaivoronski, A. A., Nesse, P. J., Erdal, O. B. Internet service provision and content services: paid peering and competition between internet providers[J]. Netnomics Economic Research & Electronic Networking, 2017, 18(1):43-79.
- [4] Moosavi, A., Sabamoniri, S., Aliakbari, M. F. MLC-MAC: A Multi-Layered Cluster-based MAC Protocol to Reduce Competition and Traffic in Wireless Sensor Networks[J]. International Journal of Computer Science and Network Security, 2017, 17(7): 247-257.
- [5] Li, F., Cao, J., Wang, X., et al. Enabling Software Defined Networking with QoS Guarantee for Cloud Applications[A].// IEEE, International Conference on Cloud Computing[C]. Honolulu: IEEE, 2017:130-137.
- [6] Kim, S., Park, S., Kim, Y., et al. VNF-EQ: dynamic placement of virtual network functions for energy efficiency and QoS guarantee in NFV[J]. Cluster Computing, 2017, 20(3):2107-2117.
- [7] Domańska, J., Domański, A., Czachórski, T. Hidden Markov Models in Long Range Dependence Traffic Modelling[A].// International Conference on Distributed Computer and Communication Networks[C]. Moscow: Springer, Cham, 2017:75-86.
- [8] Jin, X., Min, G. Modelling and Analysis of an Integrated Scheduling Scheme with Heterogeneous LRD and SRD Traffic[J]. IEEE Transactions on Wireless Communications, 2013, 12(6):2598-2607.
- [9] Feng, S., Wang, X., Sun, H., et al. A better understanding of long-range temporal

- dependence of traffic flow time series[J]. *Physica A Statistical Mechanics & Its Applications*, 2018, 492:639-650.
- [10] Li, M. Record length requirement of long-range dependent teletraffic[J]. *Physica A Statistical Mechanics & Its Applications*, 2017, 472:164-187.
- [11] Porshnev, S. V., Bozhalkin, D. A. The study of self-similarity of the traffic transmitted in the backbone Internet channel[A].// *Dynamics of Systems, Mechanisms and Machines*[C]. Omsk: IEEE, 2017:1-7.
- [12] Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., et al. K-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases[J]. *Networks & Heterogeneous Media*, 2017, 3(2):371-393.
- [13] Sarla, P., Doodipala, M. R., Dingari, M. Self Similarity Analysis of Web Users Arrival Pattern at Selected Web Centers[J]. *American Journal of Computational Mathematics*, 2016, 06(1):17-22.
- [14] Barthélemy, M., Gondran, B., Guichard, E. Large scale cross-correlations in Internet traffic[J]. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 2002, 66(2):056110.
- [15] Wu, X. F., Yu, K., Wang, X. On the growth of Internet application flows: A complex network perspective[A]. //Proc. of 30th IEEE International Conference on Computer Communications (INFOCOM) [C]. Shanghai: IEEE, 2011:2096-2104.
- [16] 王者超, 乔丽苹. 弹性力学-高等学校规划教材[M]. 中国建筑工业出版社, 2016.
- [17] 米海珍. 弹性力学(修订版)[M]. 清华大学出版社, 2016.
- [18] Lewis, T. G. *Network Science: Theory and Applications*[M]. Wiley Publishing, 2009.
- [19] Menezes, M., Kim, S., Huang, R. Constructing a Watts-Strogatz network from a small-world network with symmetric degree distribution[J]. *Plos One*, 2017, 12(6):e0179120.

- [20] Bonald, T., Massoulié, L. Impact of fairness on Internet performance[A].// Acm Sigmetrics International Conference on Measurement & Modeling of Computer Systems[C]. Cambridge: ACM, 2001:82-91.
- [21] Mozo, A., López-Presa, J. L., Anta, A. F. A distributed and quiescent max-min fair algorithm for network congestion control [J]. Expert Systems with Applications, 2018, 91:1-26.
- [22] Hu, B., Yang, Y., Wang, Z., et al. CCMF: A Congestion Control method with Fairness for multipath transport[A].// International Conference on Information Networking[C]. Da Nang: IEEE, 2017:689-694.

第四章 基于业务流突变的 服务质量波动传播特性与控制算法研究

网络环境瞬息万变，尤其是业务突发性的增强，使得业务服务质量之间的耦合作用越来越明显，并表现出明显的动力学传播性^[1-4]，业务服务质量难以有效保障。如网络热点话题在网络中的传播会瞬时占用大量网络资源，一条或几条突发业务就可能引发全网业务服务质量的波动。如何针对服务质量波动影响进行建模和分析，并提出有效的控制方案成为目前亟待解决的难题。

本章针对通信网络中突变业务流的传播动力学特性及控制机制进行研究：基于业务流服务质量之间的耦合关联关系，重点研究当网络中出现突变流时，业务流之间的竞争扩散行为与传播动力特性。具体包括，建立突变业务流的影响传播模型，研究突变流对其各阶邻居流的影响扩散行为，分析不同网络结构、不同网络状态下突变流的传播范围、传播强度等规律。此外，基于突变流影响传播模型，提出有效的突变业务流传播控制方案，保障业务服务质量。

4.1 基于耦合关联的突变业务流影响传播模型

一些学者对突变流造成的网络影响传播方面开展了相关研究^[5-7]，尤其是对突变业务流引发的网络拥塞传播现象进行了建模分析^[8-10]。Jiang^[7]等人对突变业务流及其影响进行建模，突变业务流量会造成短时间内的影响扩散，使其他业务传输时间不断增加。Ohsaki^[9]研究了网络拥塞在路由器之间的传播，发现拥塞传播速度受链路带宽和传播时延的影响，且随着网络流量突发性的减小与随机性的增加，拥塞传播周期变得不明显。Stéger 等人^[10]揭示了网络拥塞传播背后的机制，发现距离实际拥塞节点最近的路由器更易出现数据包丢失现象，而由于数据包丢失，计算机增加了数据发送速率，使得拥塞不断加剧并将拥塞的节点逐渐向下游推移。文献^[11]采用真实网络突发业务数据，提出了一种基于自回归 AR 模型的 Hurst 系数估计方法，比传统的 R/S 估计方法具有更高的估计精度，能更好地反映真实网络突发业务流量的自相似性。综合上述研究成果，目前针对突变业务流的研究主要集中在自身特性或突变业务流对网络层面的影响(路由器、链路等)，不足以揭示突变业务流对业务层面的传播影响，而业务层面服务质量的高低直接

关系到用户的网络体验。

本文基于业务层面建立了网络突变业务流的影响传播模型。为了直观地反映影响传播过程，将业务流耦合关联网络表示为具有层次化结构的网络，以初始突变业务流作为传播源，依次将其余业务流进行分类，使用“邻居阶数”描述突变业务流与其他流之间的邻近关系；并将业务服务质量间的交互作用通过“耦合概率”的方式进行表述，进而将影响从突变业务流按照某种规则扩散至各阶邻居业务。基于此传播模型，重点研究了动态网络环境下突变业务流的动力学传播特征，分析突变业务流耦合影响的全网传播机理。

4.1.1 传统传播模型

传统传播模型中应用最广泛的是传染病模型。传染病模型不仅刻画了疾病在人群中的传播机理，更应用于社会、经济、网络等众多领域。这里简单介绍几种经典的传染病模型，有助于理解网络动力学过程的基本原理。

(1) SI 模型^[12]

在典型的传染病数学模型中，疾病在寄主体内的动力学过程可以简化为少数基本疾病状态之间的转变。最简单的形式只有两个状态，易感态 (Susceptible State) 和感染态 (Infected State)，处于易感态的个体自身没有疾病，但如果和感染态的人接触，则极有可能被传染。

令 $S(t)$ 表示 t 时刻的易感人群数， $I(t)$ 表示 t 时刻的感染人群数。由于疾病的传播过程是随机的，因此感染人数并不唯一确定，即如果疾病不止一次在同一个群体进行传播，即使在相同的条件下，每次易感人群和感染人群数也可能不一样。为避免该问题，定义 S 和 I 为易感人群和感染人群的期望值。

只有感染者与易感者接触时，疾病才会传播。假设一个群体的总人数为 n ，那么随机遇到易感者的概率为 S/n ，因此，一个感染者在单位时间内平均与 $\beta S/n$ 个易感者接触。由于网络平均感染人群数目为 I ，故单位时间新增感染者的平均速率为 $\beta SI/n$ ，由此得到关于 I 变化率的微分方程：

$$\frac{dI}{dt} = \frac{\beta SI}{n} \quad (4-1)$$

同时，易感者的数量 S 以同样的速率下降：

$$\frac{dS}{dt} = -\frac{\beta SI}{n} \quad (4-2)$$

定义感染者所占的比例为 $i = I/n$ ，利用 $s = 1 - i$ 化简公式 (4-1) 得：

$$\frac{di}{dt} = \beta(1-i)i \quad (4-3)$$

联合方程 (4-1) 至 (4-3) 解得：

$$i(t) = \frac{i_0 e^{\beta t}}{1 - i_0 + i_0 e^{\beta t}} \quad (4-4)$$

以上是关于疾病传播的 SI 模型。

(2) SIS 模型^[13]

SI 模型的一个拓展就是允许再次感染的发生，即对于虽然恢复但没有出现免疫能力或免疫力很低的人，很有可能再次被感染。在此情况下，最简单的模型是 SIS 模型。模型中依然只有两个状态：易感态和感染态，但感染人群恢复后，再次回到易感态，恢复速率为 γ 。该模型的微分方程如下：

$$\begin{cases} \frac{di}{dt} = \beta s i - \gamma i \\ s + i = 1 \\ \frac{ds}{dt} = -\beta s i + \gamma i \end{cases} \quad (4-5)$$

对上式求解得：

$$i(t) = (1 - \gamma / \beta) \frac{C e^{(\beta - \gamma)t}}{1 + C e^{(\beta - \gamma)t}} \quad (4-6)$$

其中，积分常数 C 由 i 的初值决定。

(3) SIR 模型^[14]

在 SI 模型中，个体一旦被感染便永远处于感染态。而许多真实疾病的传播过程中，由于人的免疫系统会逐渐消灭病原体，经过一段时间，被感染的人会恢复健康且保持免疫力，防止了再次感染的发生。在此基础上，SIR 模型加入了节点恢复机制，即引入了节点移除 (Removed) 状态。每个个体以概率 γ 变为移除

态，不再参与感染过程。SIR 模型的数学描述为：

$$\begin{cases} \frac{di}{dt} = \beta si - \gamma i \\ \frac{ds}{dt} = -\beta si \\ \frac{dr}{dt} = \gamma i \end{cases} \quad (4-7)$$

其中， $s + r + i = 1$ 。

(4) SIRS 模型^[15]

SIRS 模型中个体从感染态恢复之后，同 SIR 模型一样获得免疫能力，但免疫能力只是暂时的，在某段时间后，该个体会失去免疫能力而重新变为易感染者。引入新参数 δ 表示个体丧失免疫能力的平均速度，SIRS 模型的方程可以表示为：

$$\begin{cases} \frac{ds}{dt} = \delta r - \beta si \\ \frac{di}{dt} = \beta si - \gamma i \\ \frac{dr}{dt} = \gamma i - \delta r \end{cases} \quad (4-8)$$

其中， $s + r + i = 1$ 。

除上述几类模型外，还提出了许多其他传染病模型，用来模拟特定类型的疾病传播。同时也引进了其他状态，例如“潜伏”态，以此来表示被感染的人还没有发展到能够传播疾病的阶段；再如易感态之前的初始免疫态，通常表示新生儿拥有的母体免疫阶段。此外，有一些模型允许新个体的加入，如新生儿或移民的出现等。

4.1.2 突变业务流影响传播模型

传统传染病模型（SI，SIS，SIR，SIRS）基于完全混合假设，即感染节点把疾病传染给任一易感态节点的概率都相等。然而这并不符合网络业务之间的关联关系，基于第三章建立的业务流耦合关联网络，只有存在连边的业务之间才会存在服务质量相互影响，且依据连边权值大小的不同，业务间的耦合影响也会不同。此外，传染病模型刻画的个体状态是离散的，忽略了个体状态在疾病过程中的连续性变化，而网络业务的传输状态通常随着网络的变化而动态连续改变。因此，

传统传染病模型尚不能准确刻画网络突变流对业务服务质量的动态影响过程，需重新建立突变业务流影响传播模型。

依据业务间共享链路情况定义突变流和网络其他流的“邻居阶数”关系，便于后文清晰描述突变流的影响传播的过程。具体而言：

一阶邻居业务：与突变流直接共享一条或几条链路的业务称作突变流的一阶邻居；

二阶邻居业务：与突变流的一阶邻居共享一条或几条链路的业务称作突变流的二阶邻居，但不包括突变流的一阶邻居；

n 阶邻居业务：与突变流的 $n-1$ 阶邻居共享一条或几条链路的业务称作突变流的 n 阶邻居，但不包括一阶至 $n-1$ 阶邻居。

依据业务间的邻居关系，分析突变业务流对外产生影响的路径特点。在网络拥塞情况下，路由器中流量都要受到排队机制、队列长度和链路带宽容量的限制，同一链路中的业务将竞争有限的网络资源。此时若突变流在很短的时间内占用大量的网络资源，与突变业务流共享路由器的一阶邻居业务会因受到竞争影响失去对网络资源（带宽）的占用。由于 TCP 反馈机制的存在，一阶邻居业务在整个传输路径上的带宽都会减少（带宽保持一致）。一阶邻居业务带宽的减少将继续对与其共享链路和路由器的二阶邻居业务产生影响。以上过程重复进行，突变流的影响最终通过各阶邻居业务间的耦合作用传播至整个网络。

为清晰反映这一动态过程，构建了具有层次化结构的业务流耦合关联网络 $G=(V,E)$ ，其中， V 表示网络中的业务流集合， E 表示业务间的连边集合。本节基于第三章建立的业务流耦合关联网络，重新构建层次化的网络模型。具体而言，将突变流置于网络中心，1 至 n 阶邻居业务依次向外延展分布，影响传播方向由突变流向其各阶邻居业务依次传播，如图 4-1（b）所示。业务之间的连边存在与否依然根据两者是否在网络中存在共享链路决定（图 4-1（a）），即只有存在共享链路的业务之间才存在连边。此外，重新定义了业务流耦合关联网络的边权矩阵 $W(G)=(w_{i,j})_{N \times N}$ 。边权 $w_{i,j}$ 表示业务间共享链路数量，用以衡量业务间相互影响的强度。以图 4-1 为例，业务 5 和业务 7 同时共享两条链路，因此 $w_{5,7} = w_{7,5} = 2$ 。

首先，建立任意相邻邻居业务间的传播模型。 $F=\{F_1, F_2, \dots, F_i, F_{i+1}, \dots\}$ 表示网络中所有的业务流集合，其中， F_i 为第 i 阶邻居业务的集合。假设突变业务流的

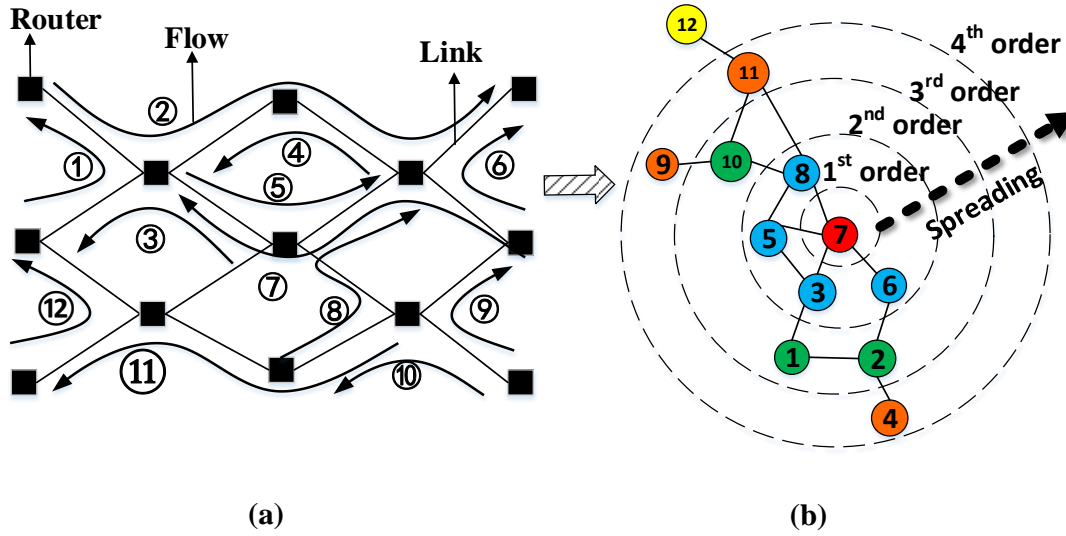


图 4-1 (a) 网络结构与业务分布; (b) 层次化的业务流耦合关联网络; 其中处于网络中心处的业务流 7 表示突变流, 向外依次是它的 1 至 4 阶邻居。

影响已传递至其第 i 阶邻居, 进一步刻画影响由第 i 阶传递至第 $i+1$ 阶的过程。

$$\begin{cases} F_i = \{f_i^1, f_i^2, \dots, f_i^j, \dots, f_i^N\} \\ F_{i+1} = \{f_{i+1}^1, f_{i+1}^2, \dots, f_{i+1}^k, \dots, f_{i+1}^M\} \end{cases} \quad (4-9)$$

其中, N 表示集合 F_i 的业务个数, 即第 i 阶邻居数目。 M 表示集合 F_{i+1} 的业务个数, 即第 $i+1$ 阶邻居数目。

对于第 i 阶邻居中的第 j 条业务 f_i^j (图 4-2), 假设其带宽受到的影响为 ΔB_i^j 。 $\Delta B_i^j = B_i'^j - B_i^j$, B_i^j 为初始带宽, $B_i'^j$ 为其受到关联影响后的带宽。 $\Delta B_i^j \leq C$, C 为链路总容量。 L_i^j 为业务 f_i^j 的路径长度。 F_{i+1}^j ($F_{i+1}^j \in F_{i+1}$) 表示第 $i+1$ 阶邻居业务中与业务 f_i^j 存在共享链路的业务集合, 即在业务流耦合关联网络中与 f_i^j 存在连边的 $i+1$ 阶邻居。考虑到不同链路中的拥塞情况不尽相同, f_i^j 对下一阶邻居业务的影响也将不同, 因此, 将集合 F_{i+1}^j 按照 f_i^j 的传播路径进行分类:

$$F_{i+1}^j = \{F_{i+1}^{j(1)}, F_{i+1}^{j(2)}, \dots, F_{i+1}^{j(q)}, \dots, F_{i+1}^{j(L_i^j)}\} \quad (4-10)$$

式 (4-10) 中子集元素表示为图 4-2 中不同虚线圈内的业务。

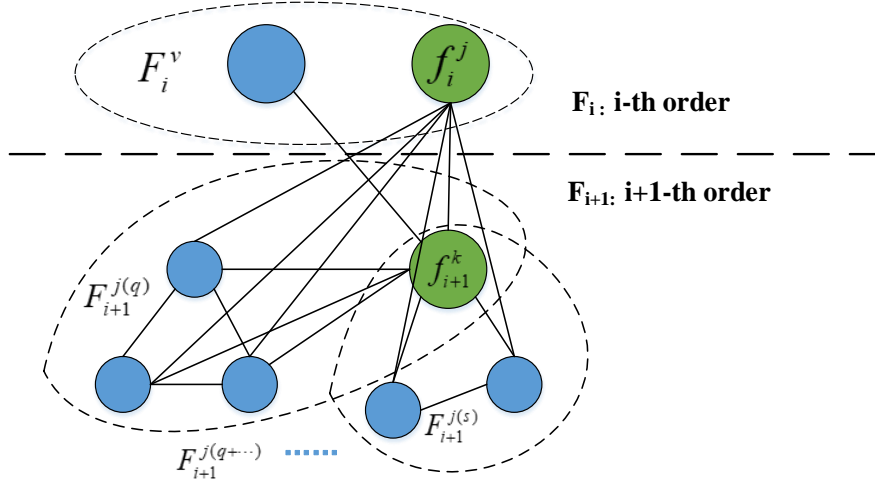


图 4-2 第 i 阶业务 f_i^j 对第 $i+1$ 阶业务 f_{i+1}^k 的影响示意图。每个虚线圈内的业务均经过同一链路，业务之间全连接。

另外，由于每一条业务都同时经过多条链路，因此 F_{i+1}^j 的子集间将存在重复元素。例如，如果业务 f_{i+1}^k ($f_{i+1}^k \in F_{i+1}^j$) 流经业务 f_i^j 的第 $\{q, q+1, \dots, s\}$ 链路，那么 f_{i+1}^k 将同时属于 $f_{i+1}^k \in F_{i+1}^{j(q)}, F_{i+1}^{j(q+1)}, \dots, F_{i+1}^{j(s)}$ ，如图 4-2 所示，不同虚线圈之间存在交叉元素 f_{i+1}^k 。

在 f_i^j 的第 q 条链路中，第 $i+1$ 阶业务 f_{i+1}^k 受到第 i 阶业务 f_i^j 的带宽影响为：

$$\Delta B_{i+1}^{k(q)} = p_{j,k} \frac{-\Delta B_i^j}{Z_q} \frac{1}{\varphi_{i+1}^k \sum_{\alpha=1} \frac{1}{\varphi_{i+1}^\alpha}} \quad (4-11)$$

其中， Z_q 为集合 $F_{i+1}^{j(q)}$ 中的元素个数，即业务 f_i^j 经过的第 q 条链路中其他业务的数目。假设网络中的业务具有优先级 $\Gamma = \{\varphi\}$ ，不同优先级的业务带宽可被不同程度的挤压，优先级与带宽受挤压程度成反比。 φ_{i+1}^k 表示业务 f_{i+1}^k 的优先级。 $p_{j,k}$ 表示业务 f_i^j 对业务 f_{i+1}^k 的影响概率，称作“耦合概率”，表示为：

$$p_{j,k} = \frac{w_{j,k}}{L_i^j} \quad (4-12)$$

其中, $w_{j,k}$ 表示业务 f_i^j 和业务 f_{i+1}^k 的共享链路数目。如果两个业务之间有更多的共享链路, 则它们之间相互影响的概率越大。

根据业务在其传输路径中带宽的一致性, 业务 f_i^j 对业务 f_{i+1}^k 造成的带宽变化取决于影响最大的链路:

$$\Delta B_{i+1}^k(j) = \max(\Delta B_{i+1}^{k(q)}, \Delta B_{i+1}^{k(q+1)}, \dots, \Delta B_{i+1}^{k(s)}) \quad (4-13)$$

通常情况下, 业务 f_{i+1}^k 会受到来自第 i 阶邻居中不止一条业务的影响, 假设 F_i^v 为第 i 阶邻居中与业务 f_{i+1}^k 存在直接影响关系的业务集合, 如图 4-2 所示。最终业务 f_{i+1}^k 受到第 i 阶业务影响的总和为:

$$\Delta B_{i+1}^k = \sum_{j \in F_i^v} \Delta B_{i+1}^k(j) \quad (4-14)$$

网络通常使用丢包率、吞吐量、往返时间、队列延迟等参数刻画业务传输性能^[16-17]。基于此, 本文使用业务带宽变化百分比, 即业务丢包率的变化, 衡量突变业务流对网络业务的影响。

B_{i+1}^k 为业务 f_{i+1}^k 的初始传输带宽, 初始带宽分布利用第三章建立的业务流弹性物理模型求解, 那么业务 f_{i+1}^k 的带宽抖动为:

$$\lambda_{i+1}^k = \frac{\Delta B_{i+1}^k}{B_{i+1}^k} \quad (4-15)$$

λ 反映了突变业务流造成的网络业务服务质量抖动情况。 $\lambda < 0$ 时, 表明业务带宽受到突变流影响后减小; $\lambda > 0$ 时, 表明业务带宽受到突变流影响后增加; $\lambda = 0$ 时, 表明业务带宽不受影响。 λ 绝对值越高, 网络业务受到突变流的影响越大, 也表明突变流的影响强度越强。

至此, 任意两阶相邻业务之间的影响传播建模完毕, 通过逐层传播, 可得到整个突变业务流的影响传播过程。令 $\bar{\lambda}$ 表示网络中业务带宽平均受突变流影响的程度:

$$\bar{\lambda} = \langle |\lambda_f| \rangle_{f \in F} \quad (4-16)$$

其中, F 表示网络中所有的业务流集合。

4.1.3 模型精确度

(1) 耦合概率验证

在传播模型中，定义业务耦合概率与业务之间共享链路的数量成正比 $p \propto w$ （公式 4-12），即两个业务流之间有更多的共享链路将增加它们在链路中相互影响的概率，业务带宽更有可能受到影响。为验证该定义的合理性，利用网络仿真软件模拟了业务在动态网络环境下带宽的变化情况。

图 4-3 是美国 USNET 的拓扑结构（24 个节点和 43 条链路）。假设链路带宽容量为 $C=10$ （Gbps）。200 条业务流在网络中随机选择起点和终点，并通过最短路径算法进行选路。业务初始发送速率满足正态分布 $D \sim N(5, 0.6)$ ，业务优先级随机设定为 $\varphi=1, 2, 3$ ，业务带宽分配满足比例优先级带宽算法。

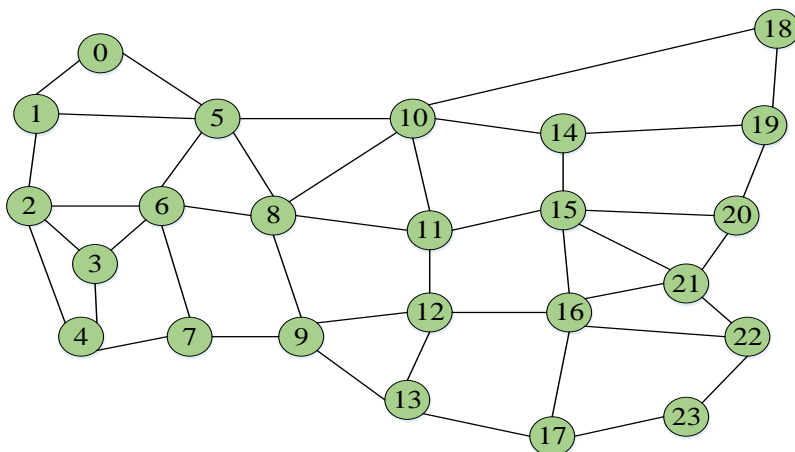


图 4-3 USNET 网络结构（24 个节点 43 条链路）

不失一般性，仿真实验中随机选择一条路径长度为 4 的业务作为突变业务流，将其带宽增加 5 倍，并同时关注突变业务流一阶邻居的带宽变化情况。实验结果如图 4-4 所示，x 轴表示一阶邻居业务的带宽变化百分比，y 轴表示突变业务流和一阶邻居之间的平均共享链路数 w 。结果显示，一阶邻居中受影响较大的业务通常是与突变业务流有更多共享链路数的业务，如一阶邻居中带宽受影响程度大于 80% 的业务与突变业务流的共享链路数通常大于 3，即耦合概率大于 0.75。因此，可以得出结论：两个业务流之间的共享链路越多，相互影响的概率越大，验证了耦合概率定义的合理性。

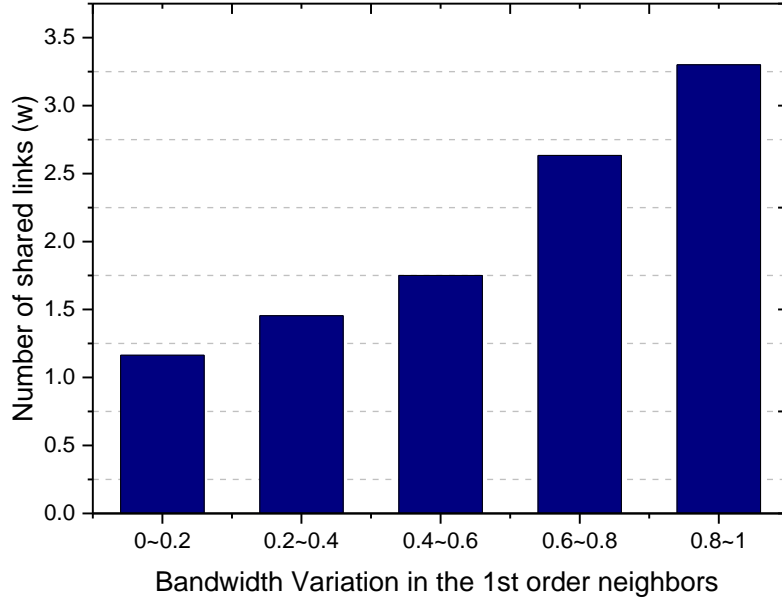


图 4-4 业务带宽受影响程度与共享链路数关系

(2) 模型偏差验证

定义模型偏差 δ :

$$\delta = \frac{\sum_{i=1}^N \left| \frac{\Delta B_{model}^i - \Delta B_{actual}^i}{\Delta B_{actual}^i} \right|}{N} \quad (4-17)$$

其中, N 表示网络中业务流总数。 ΔB_{model}^i 表示通过模型仿真得出的突变业务流引起的网络业务平均带宽变化值, ΔB_{actual}^i 表示通过网络仿真软件模拟得到的实际业务流平均带宽变化。

实验使用 USNET (图 4-3) 作为基础网络, 链路带宽容量为 $C = 10$ (Gbps)。业务流在网络中随机选择起点和终点, 初始发送速率 (或业务初始需求带宽) 满足正态分布 $D \sim N(5, 0.6)$, 业务优先级随机设定为 $\varphi = 1, 2, 3$, 业务带宽分配满足比例优先级带宽算法。网络中的业务数目逐渐从 50 增加到 240 以增加网络拥塞链路所占的比例。突变业务流随机选择, 并将其带宽增加至初始 5 倍值。

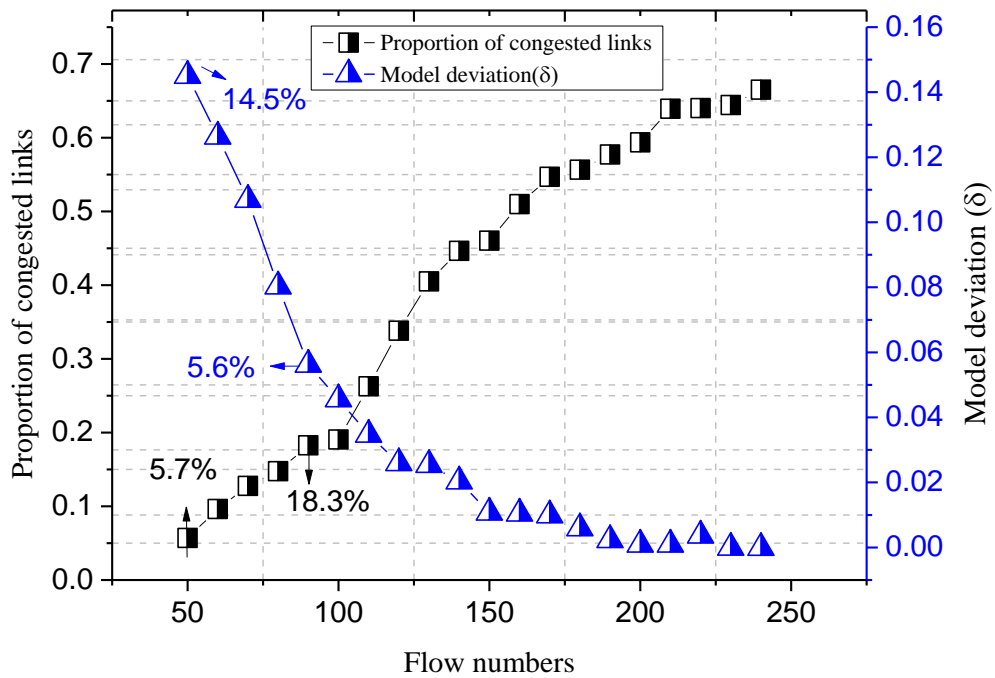


图 4-5 模型偏差。横坐标表示网络中业务流的数目，左侧纵坐标表示网络中 congested 链路的比例，右侧纵坐标表示模型的偏差百分比。

从图 4-5 可以看出，随着业务数目的增长，网络中 congested 链路的比例迅速增加，但模型的偏差值急剧下降，即网络越 congested，模型精确度越高。当网络中 congested 链路比例为 5.7% 时，模型偏差为 14.5%，但当网络中拥有 18.3% 以上的 congested 链路时，模型偏差已经小于 5.6%。

近年来，移动数据爆发式增长，严重加剧了网络 congested。我国工业和信息化部发布的“2017 年通信业统计公报”显示，2017 年我国移动互联网数据流量已达 246 亿 GB，且从 2012-2017 年一直呈指数增长（图 1-3）。业务量级的爆发式增长使得网络 congested 将在未来的通信网络中更为常见。因此，本节提出的模型适用于目前大多数的网络 congested 情况。

4.1.4 实验仿真与结果分析

利用突变业务流影响传播模型，本节基于动态网络环境：不同突变强度、不同网络 congested 程度及不同网络结构类型，分析突变业务流的传播特性。

(1) 不同强度突变流

各阶邻居业务带宽受影响程度通过平均带宽变化率反映，计算公式如下：

$$\lambda_i = \left\langle \frac{\Delta B_{f_i}}{B_{f_i}} \right\rangle \quad (4-18)$$

其中， $f_i \in F_i$ ， F_i 为第 i 阶邻居业务的集合。 λ_i 体现了突变业务流对其第 i 阶邻居业务的影响强度。

实验环境设置如下：在 USNET 网络（图 4-3）中 200 条业务随机选择起点和终点，业务按照最短路径算法 Dijkstra 选择路径。网络的初始链路容量为 5（Gbps），业务带宽需求满足正态分布 $D \sim (2, 0.4)$ ，业务优先级随机设定为 $\varphi=1, 2, 3$ ，业务需求与链路容量可相比使得网络有相当部分的链路处于拥塞状态。依据第 3.1 节建立的弹性物理模型，可以得到业务初始的带宽分配情况 B_{f_i} 。随机选择一条业务流作为突变业务流，并多次、逐步突增其带宽，依据第 4.1.2 节建立的突变流影响传播模型，可以得到业务带宽的动态变化值 ΔB_{f_i} 。

为反映突变业务流的影响传播特性并保持图的清晰度，图 4-6（a）给出突变业务流对前三阶邻居业务的影响结果。从图 4-6（a）可以看出，随着突变业务流的带宽增长，各阶邻居业务平均带宽受影响程度均逐渐增加（绝对值），但对相邻邻居业务的影响相反：奇数阶邻居的带宽变化为负数，即业务流带宽减小；偶数阶邻居的带宽变化为正数，即业务流带宽增加。在网络拥塞、链路饱和的情况下，当某一业务增加（或减少）带宽时，由于链路总容量有限，与其共享链路的其他业务将减少（或增加）带宽，带宽增减趋势恰好相反，进而出现了突变流对各阶业务影响的正负交替结果。

此外，随着邻居阶数的增长，突变业务流的影响强度急剧下降。在图 4-6（a）中，突变流影响强度的平均下降程度为：从一阶邻居到二阶邻居，下降 22 倍左右；从二阶邻居到三阶邻居，下降 50 倍左右。根据公式（4-11），影响强度的衰减趋势与每阶邻居中的业务数目相关。图 4-6（b）显示了不同邻居阶数的业务数量，业务数目先增后减。综上所述，在网络拥塞环境下，突变业务流的影响传播是一个随着邻居阶数增长而大幅下降的过程，且对奇数阶邻居和偶数阶邻居的影响正负交替进行。

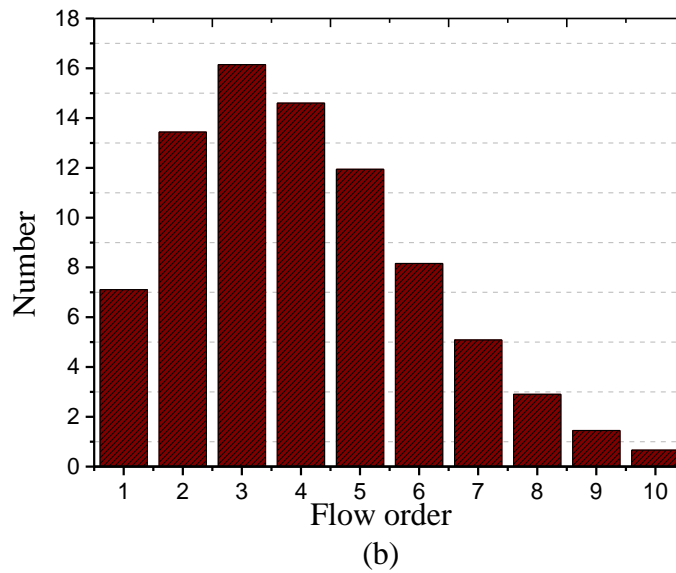
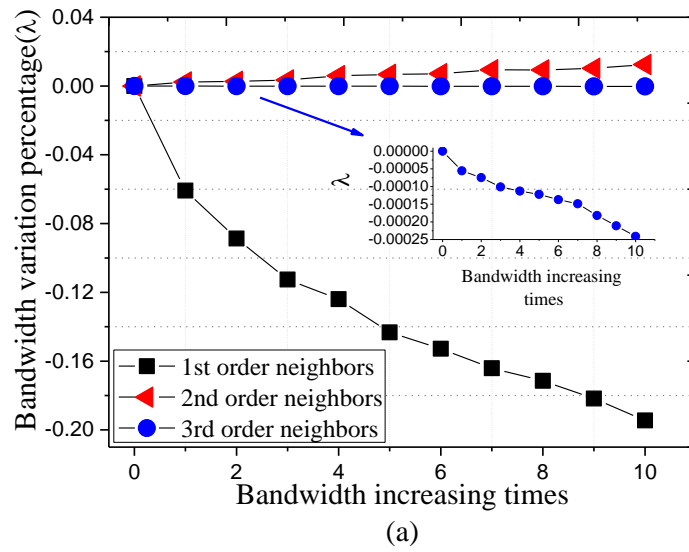


图 4-6 (a) 突变流对其前三阶邻居业务的影响强度变化；(b) 突变流各阶邻居数目。横坐标为邻居阶数，纵坐标为数目。

本文搭建了实际网络环境对图 4-6 中模型仿真结果的可靠性进行验证。PTN 网络 (Packet Transportation Network) 是基于 PW 伪线技术和 LSP 标签技术的分组业务传送网络, 利用 4 台 PTN 设备搭建了如图 4-7 所示的网络结构, 接入两台 IP 流量分析仪 (MD1230B) 注入业务并对传输性能参数进行测量。

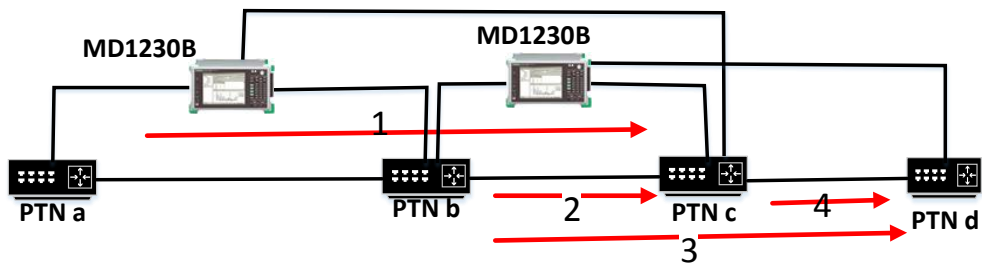


图 4-7 PTN 测试网络

表 4-1 业务参数设置

Service	Priority	Rate	PW (CIR/PIR)	PW	LSP	LSP (CIR/PIR)
1	AF	40M	40/60M	1	1	100M /100M
2	BE	30M	0/30M	2	2	100M /100M
3	BE	30M	0/30M	3	2	100M /100M
4	BE	70M	0/70M	4	3	100M /100M

业务参数设置如表 4-1 所示。业务 1 为高优先级业务（AF：Assured Forwarding），业务 2-4 为低优先级业务（BE：Best Effort）。测试前 10s 时，业务 2-4 自由传输，后 10s 开始注入突变业务 1 并不断增加其带宽（40-45-50-55-60）。通过 IP 流量分析仪得到业务 2-4 在 10s~20s 内的丢包率变化情况，如图 4-8 所示。可以看出，1) 突变流对奇偶阶邻居业务的影响相反，对奇数阶业务（业务 2，3）产生负面影响，对偶数阶业务（业务 4）产生正面影响；2) 突变流的影响强度随着邻居阶数增长明显下降。测试结果与模型仿真结果（图 4-6）一致，验证了突变流影响传播模型的正确性。

（2）不同网络拥塞程度

网络拥塞是导致业务服务质量间相互影响的直接原因，因此，本文继续评估网络拥塞程度对突变业务流影响传播的影响。定义网络拥塞系数 Q 描述网络的平均拥塞程度：

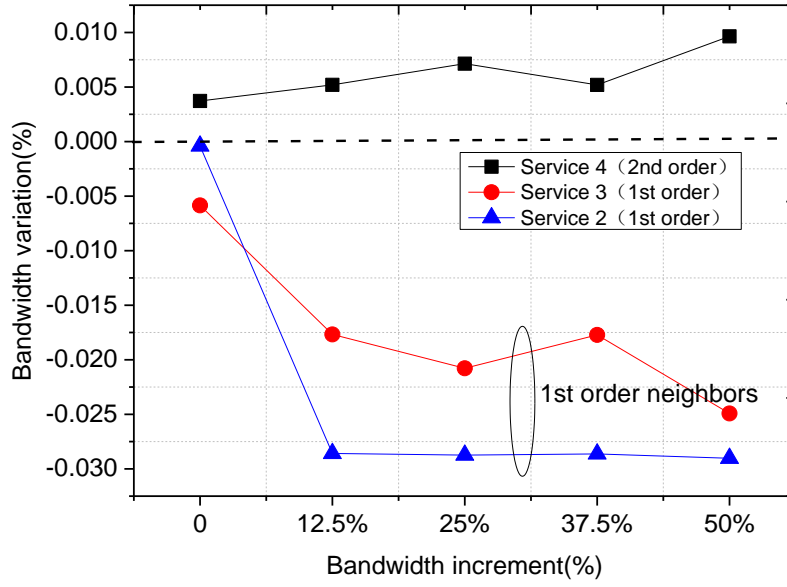


图 4-8 PTN 测试网络中突变流的影响传播

$$Q = \left\langle \frac{D_l - C}{C} \right\rangle \quad (4-19)$$

其中, $l \in L$, L 表示网络链路集合。 D_l 是链路 l 中业务带宽需求总和, C 是链路带宽容量。拥塞系数 Q 可以通过改变链路容量 C 进行调整。

实验使用 USNET 作为基础通信网络, 200 条业务随机选择起始节点和目的节点, 并按照最短路径算法选择路径。网络的初始链路容量为 5Gbps (可调节), 业务带宽需求满足正态分布 $D \sim N(2, 0.4)$, 业务优先级随机设定为 $\varphi = 1, 2, 3$ 。从网络中所有的业务流中随机选取一条作为突变业务流, 将其带宽增加为原来的 5 倍。图 4-9 为突变业务流影响强度随网络拥塞系数的变化。从图中可以看出, 突变业务流的影响强度明显受到网络拥塞程度的影响。当拥塞系数逐次增加 10% 时, 突变业务流对其前三阶业务流的影响产生以下变化: 对一阶业务的影响平均增加 34.38%, 对二阶业务的影响平均增加 3.32%, 对三阶业务的影响平均增加 0.33%, 网络拥塞程度的改变主要影响了突变流对其一阶邻居的影响强度。

结合 (1) 和 (2) 中得出的主要结论, 突变业务流影响传播的基本特征如下: 在网络拥塞环境下, 突变业务流对其奇数阶邻居会产生负面影响 (带宽减少), 对偶数阶邻居产生正面影响 (带宽增加), 并且随着邻居阶数的增加, 其影响强

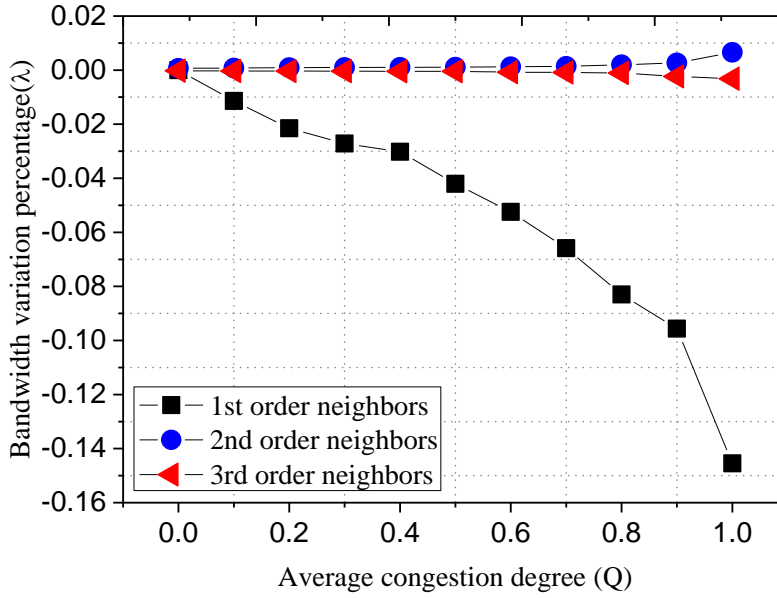


图 4-9 突变流影响强度随网络拥塞系数的变化

度显著下降。另外，随着网络拥塞程度的加剧，突变流的影响强度也将增加，但主要影响其对一阶邻居业务的影响。

(3) 不同网络规模与结构

研究表明，网络业务流在不同类型的网络中具有不同的动态特性^[16,18]。网络结构会影响业务路径的选择、业务的整体分布以及资源的利用情况等。因此，在仿真实验中，考虑三种类型的网络：ER 随机网络，BA 无标度网络和 WS 小世界网络，并以全网业务平均带宽变化百分比 $\bar{\lambda}$ 反映突变业务流的平均影响强度。对于 ER，BA，WS 三种类型的网络，保持网络中业务总量 1000 不变。1000 条业务流随机选择起始点和终点，按照最短路径算法选择路径。为避免优先级的干扰，设置业务具有相同优先级。网络链路容量均设为 5Gbps，业务带宽需求满足正态分布 $D \sim N(2, 0.4)$ 。随机选择一条业务流作为突变业务流，并将其带宽增加 5 倍。

首先，分析网络规模对突变业务流传播强度的影响。网络规模从 20 增长至 110，网络节点数每次增加 10。如图 4-10 (a) 所示，随着网络规模增大，突变业务流的平均影响强度 $\bar{\lambda}$ 逐渐下降。以 WS 小世界网络为例，网络每增加 10 个节点，突变业务流的影响强度平均下降 14.78%。依据 4.1.2 节建立的突变业务流

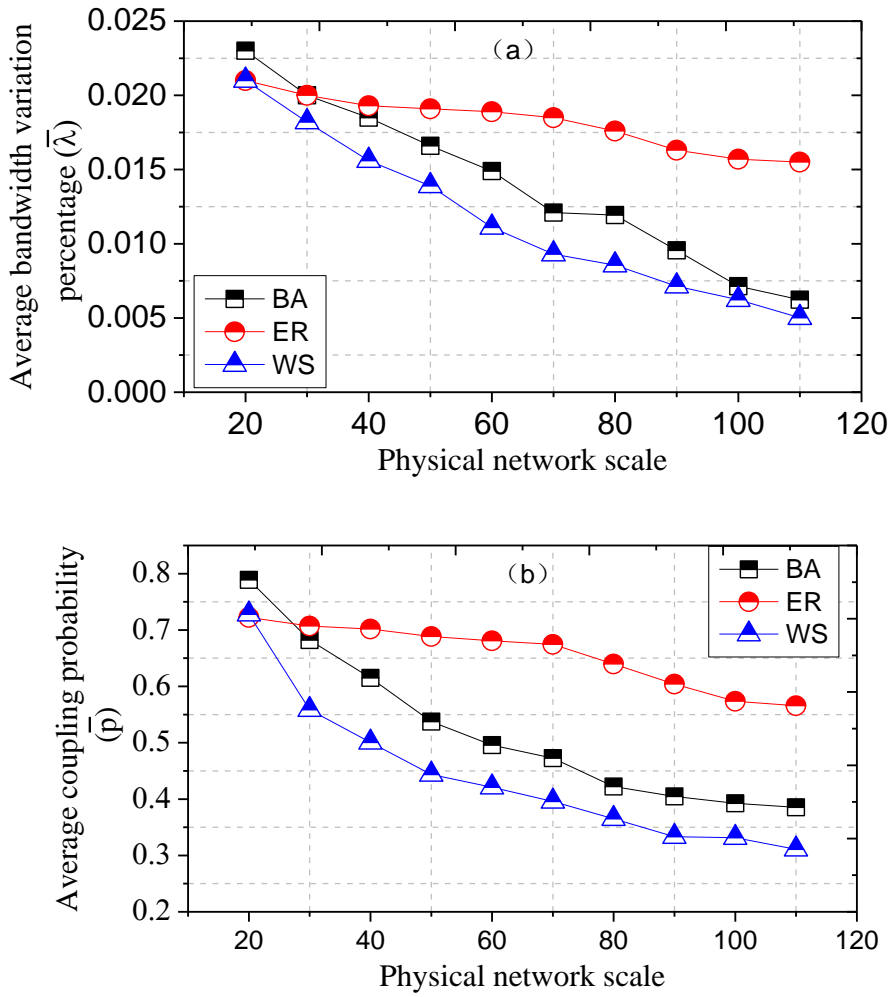


图 4-10 (a) 影响强度；(b) 耦合概率

传播影响模型可知，业务之间的耦合概率 \bar{p} 会影响突变业务流的传播强度。在传播过程中，如果业务之间的耦合概率 \bar{p} 越高，则突变业务流的影响将更容易传播至下一阶邻居（公式 4-11）。从图 4-10 (b) 可以看出业务之间耦合概率 \bar{p} 随着网络规模的不断增长不断下降。以 WS 小世界网络为例，网络规模每增加 10 个节点，业务之间的耦合概率平均下降 8.8%。因此，耦合概率 \bar{p} 的下降最终导致了突变业务流的平均影响强度 $\bar{\lambda}$ 的下降。

其次，分析网络拓扑特征对突变业务流传播强度的影响。如图 4-10 (a) 所示，随着网络规模的扩大，WS 网络比 BA 网络和 ER 网络在降低突变业务流影响强度方面表现更为优异。WS 网络中突变业务流的平均影响强度比 ER 网络低 38.18%，比 BA 网络低 18.4%。为理解这一现象，依然从业务流耦合概率的角

度进行分析。如图 4-10 (b) 所示, 可以看出在相同物理网络规模和相同业务总数的情况下, WS 网络中业务间相互影响的耦合概率比 ER 网络和 BA 网络平均低 33.89% 和 15.93%。ER 网络的节点度服从泊松分布, 结构差异化不明显, 业务分布均匀, 使得业务间平均耦合概率最高; BA 网络中由于存在“枢纽”节点, 使得业务在“枢纽”节点附近的聚集程度较高, 因而耦合概率相较于 WS 网络略高。耦合概率 \bar{p} 的差异使得 WS 网络中突变流传播强度低于其他两种类型网络。

通过上述分析, 得出结论: 1) 当网络中业务数目固定时, 网络规模的扩大可以降低业务之间的耦合概率, 从而降低突变业务流对全网业务的平均影响强度; 2) 在相同网络规模、固定业务数目和随机业务分布情况下, 小世界类型网络中由于业务之间的耦合概率较低因而在降低突变业务流的影响强度方面优于 ER 和 BA 网络。

4.2 基于社团结构的突变业务流免疫控制算法

网络流量具有难以预测的突发性特征, 突变业务流在短时间内迅速变化, 造成流量之间激烈的带宽竞争。4.1 节建立了突变业务流的影响传播模型, 突变流引起的瞬时拥塞会导致其各阶邻居业务产生服务质量上的波动变化, 降低网络性能。本节进一步深入研究突变业务流影响传播控制方案, 对提高网络稳定性和可靠性具有重要意义。

为有效缓解突变业务流的影响, 一些学者尝试从源节点处减弱突变业务流的强度^[19-21]。Park 等人^[19]提出了一种新的流量生成算法“window based segment spacing”。该算法控制数据均匀分割成多个数据块, 将流量突变造成的影响从源头得到平缓。Torabzadeh 等人^[20]发现, 通过为突发业务分组分配不同的时间延迟, 流量的突发特性将被破坏。张鹤颖等人^[22]通过仿真实验比较了不同 AQM 机制在突变业务流环境下的丢包率。通过对 AQM 机制的控制系统分析, 发现 AQM 机制通过使报文丢弃率的改变滞后于队列长度的改变, 避免了突变业务环境下的瞬时大量丢包现象。以上方法通过增加开销的方式削弱业务流源头的突变强度, 在抑制突变业务流影响的同时降低了业务的传输效率。考虑到网络不能时刻对突变业务流量做出快速响应, 一旦检测不到突变业务流或未能采取及时措施, 波动影响将很快蔓延到整个网络。因此, 本文从另一角度出发, 提出一种通过阻断传播路径实现影响控制的方案, 使突变流影响传播控制系统更为完备。

传染病网络有许多可以借鉴的传播路径控制方法。在传染病网络中, 及时识

别病毒和快速免疫是防止流行病蔓延的关键因素^[23-25]。“免疫（Immunization）”有两方面的含义：一是通过切断网络节点与病毒节点之间的传播途径，对健康节点进行隔离；二是对节点进行预处理，使其即使接触到病毒也不会被感染，比如给网络中的计算机安装防火墙技术。此外，网络结构也会对传播行为造成影响。一些学者通过对具有社团特性的网络进行分析，发现病毒的传播效率很大程度上依赖于社团化程度的高低。Wu 等人研究了具有无标度特性和可调簇系数的社团网络^[26]，发现当社团模块度上升时，病毒传播的效率会降低，即社团结构的存在会降低病毒爆发的危险性。Salathé 等人发现社团结构特征对免疫效率有着明显的影响^[27]，免疫连接社团之间的节点比免疫度大的节点更有效。

受此启发，本节提出了一种基于业务免疫和业务社团结构的突变业务流影响传播控制算法：利用层次聚类算法，将业务划分为不同等级的社团，通过免疫不同等级社团边界的关键业务，阻断社团间的传播途径，最终将突变流的影响控制在极小范围之内。

4.2.1 算法设计

在传染病传播模型中，通过将易感染者进行免疫可以阻断疾病的传播路径。借鉴传染病案例中的免疫过程，本文提出基于社团结构的突变业务流免疫控制算法。网络中突变业务流的影响传播过程本质上是由弹性业务带宽波动引起的，因此，为抑制带宽波动传播，将一些业务的优先级设置为最高级别，保障其刚性传输（带宽不因网络动荡而发生变化），则突变业务流的影响传播路径将被阻断，本文将这一过程称之为**网络业务“免疫”**。

选择性地免疫可显著降低算法复杂度和提升算法效率。研究发现真实世界的网络大多具有一个共同的特征：社团结构。具有社团结构的网络有着独特的组织特性：社团内节点的连接比较稠密，社团间节点的连接比较稀疏。连接两个社团之间的节点和链路具有高介数中心性，而通常高介数中心性的节点或边由于控制着其他节点间消息的传递，在网络中有着很强的影响力。因此，社团结构对网络中的传播现象和控制策略有很重要的影响^[28-31]。本节将利用业务的社团结构，对社团边界业务流进行免疫，有效控制社团间的影响传播。

目前，社团划分算法主要有 GN 算法、Fast Newman 算法、BGLL 算法等。GN 算法是由 Girvan 和 Newman 提出的一种基于边介数的划分算法^[32]，具有非常高的划分精确度，但算法复杂度高，且需要事先确定划分社团的个数；Fast Newman 算法属于凝聚算法，适用于大规模稀疏网络，但算法精确度略低^[33]；

BGLL^[34]算法是一种能够用于加权网络的层次化社团结构凝聚算法，算法过程简单直观、容易实现、复杂度低且效果好。考虑到业务流耦合关联网络规模巨大，且业务间连边密集、带有权值，本文将采用 BGLL 算法。

针对本节提出的应用场景，BGLL 算法具体实现如下：

Step1: 将业务流耦合关联网络中每个业务节点看作一个社团，按一定次序依次遍历每一个节点，即对每个节点 i ，考虑将其移至邻居节点 j 的社团 c_j 之后模块度的变化 ΔQ 。如果 $\Delta Q > 0$ ，将节点 i 移至使得 ΔQ 变化最大的节点的社团中；否则，节点 i 保持不动。重复这个过程，直到任何节点的移动都不能使模块度增大。

模块度的计算方法如下：

$$Q = \frac{1}{2m} \sum_{i,j} [W_{ij} - \frac{K_i K_j}{2m}] \delta(c_i, c_j) \quad (4-20)$$

其中 W_{ij} 表示节点 i 与节点 j 之间的权值， K_i 表示网络中与节点 i 相连接的边的权重之和， c_i 表示节点 i 所属的社团，如果节点 i 与节点 j 在同一个社团，则 $\delta(c_i, c_j) = 1$ ，否则 $\delta(c_i, c_j) = 0$ 。 m 为整个网络的边权重之和。

$\Delta Q = (\text{节点 } i \text{ 移除后原社团的模块度} - \text{节点 } i \text{ 移除前原社区的模块度}) + (\text{节点 } i \text{ 移动到社团 } c_j \text{ 后社团 } c_j \text{ 的模块度} - \text{社团 } c_j \text{ 原本的模块度})$ 。

Step2: 将 Step1 得到的每一个社团看作一个新的节点，开始新一轮迭代，直到模块度不再变化。新节点之间的边权重为两个新节点（其实是两个社团）原本的权重之和。处在同一个社团中的节点之间的边导致新网络中该新节点有自环边，对构建的新网络使用 Step1 中的方法进行迭代，当网络社团结构不再改变，即出现了最大模块度值的时候停止迭代。

以上，为 BGLL 算法的核心步骤。为更精确地得到突变业务流所在的社团，并有效地进行免疫控制，两次使用 BGLL 算法，将业务流耦合关联网络划分为一级社团和二级社团。以图 4-11 为例，(a) 是利用 BGLL 算法分割得到一级社团的过程，即业务流耦合关联网络的初次社团划分；(b) 是将某一级社团继续使用 BGLL 算法得到二级社团的过程。

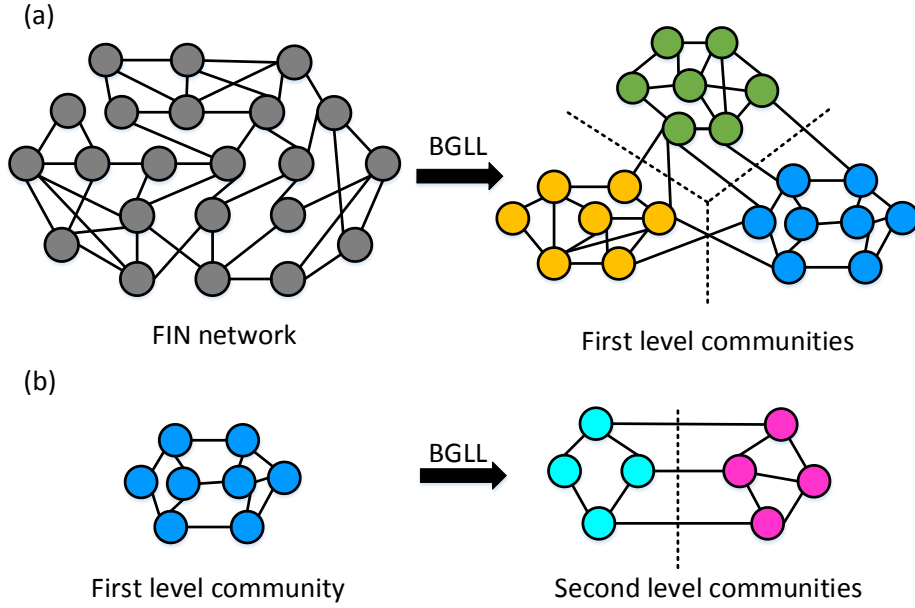


图 4-11 业务流耦合关联网络的两级社团结构。(a) 一级社团划分过程，(b) 二级社团划分过程。

基于业务流耦合关联网络的两级社团结构，相应的社团免疫控制过程也分为两个阶段。

第一阶段（phase I） 将业务流耦合关联网络按照 BGLL 算法进行第一次社团划分，分离一级社团（突变业务流属于其中的某一社团）。将一级社团的**社团边界业务流**部分免疫，切断突变业务流所在社团向外传播的路径。 k_i 为业务*i*的度数：

$$k_i = k^{in} + k^{out} \quad (4-21)$$

其中， k^{in} 表示业务*i*的一阶邻居中属于同一社团的业务个数， k^{out} 表示业务*i*的一阶邻居中属于不同社团的业务个数。为有效控制社团间的影响扩散，优先免疫 k^{out} 值较高的社团边界业务。如果多个业务具有相同 k^{out} 值，则选择具有较高 k^{in} 值的业务。

定义免疫步长。由于社团边界业务数目众多，需要有选择性的免疫，在每个控制阶段（phase I 和 phase II），业务免疫过程都分步进行，每次只免疫 X 条业务，直至一级社团之间的每条连边两侧的业务至少有一条业务已经进行了免疫，第一阶段结束。

第二阶段 (phase II) 把突变业务流所在的一级社团看作独立的网络, 利用 BGLL 算法进行社团的二次划分, 将划分得到的二级社团的社团边界业务继续进行免疫, 以进一步缩小影响传播范围。选择免疫的二级社团边界业务依然按照“先 k^{out} 后 k^{in} ”的规则进行, 直至该二级社团之间的每条连边两侧的业务至少有一条业务已经进行了免疫, 第二阶段结束。

值得注意的是, 每免疫一条业务, 在下次选择新的免疫业务时都需要重新计算业务的度数。若某条边两侧的业务中已有一条业务被免疫, 影响将不能再通过这条连边进行传播。因此, 为提高控制效率, 每免疫一条业务, 与该业务相连的连边将在下一次计算业务度数时将不再考虑。

为更清晰的反映免疫过程, 以图 4-12 所示的业务流耦合关联网络为例, 免疫过程具体如下: 图 4-12 是由 22 条业务组成的网络, 业务 1 为突变业务。第一阶段 (Phase I) 利用 BGLL 算法将网络划分为 3 个一级社团, 并按照“先 k^{out} 后 k^{in} ”的规则对社团边界业务流进行免疫。Step 1-5 分别免疫了业务 2、3、4、5、6。每个 Step 中, 只选择一条业务进行免疫 (免疫步长 $X=1$)。通过免疫了 22.7% 的业务, 一级社团间的影响传播就停止了。未受影响的业务比例为 68.2%, 这意味着 45.5% 的业务在免疫控制过程中得到额外的保护。

第二阶段 (Phase II) 将突变业务 1 所在的一级社团看作独立网络, 进行二级社团划分, 得到两个小的社团。继续按照“先 k^{out} 后 k^{in} ”的规则, 对二级社团边界业务进行免疫。通过 Step1 和 Step2 免疫业务 7 和 8, 最终将突变业务 1 的影响控制在极小的范围之内。在真实网络中, 突变业务流同时发生的数目远远超过一个。在此情况下, 第一阶段的免疫过程不需改变, 只需将第二阶段的多个子社团同时进行免疫, 本节提出的控制策略依然可以有效控制一级社团和二级社团间的影响传播, 避免多重突变流造成的叠加影响。

4.2.2 实验仿真与结果分析

在突变业务流影响传播模型中, 业务由节点表示, 业务之间的相互影响由边表示。基于该业务流耦合关联网络, 本节提出的社团免疫控制算法可以有效的控制社团间的影响传播。为衡量社团免疫控制策略的效果, $\bar{\lambda}$ 表示业务突变前后全网业务带宽平均受影响程度; $\Phi_{\bar{\lambda}=0}$ 表示业务突变前后网络中带宽未受任何影响的业务比例。

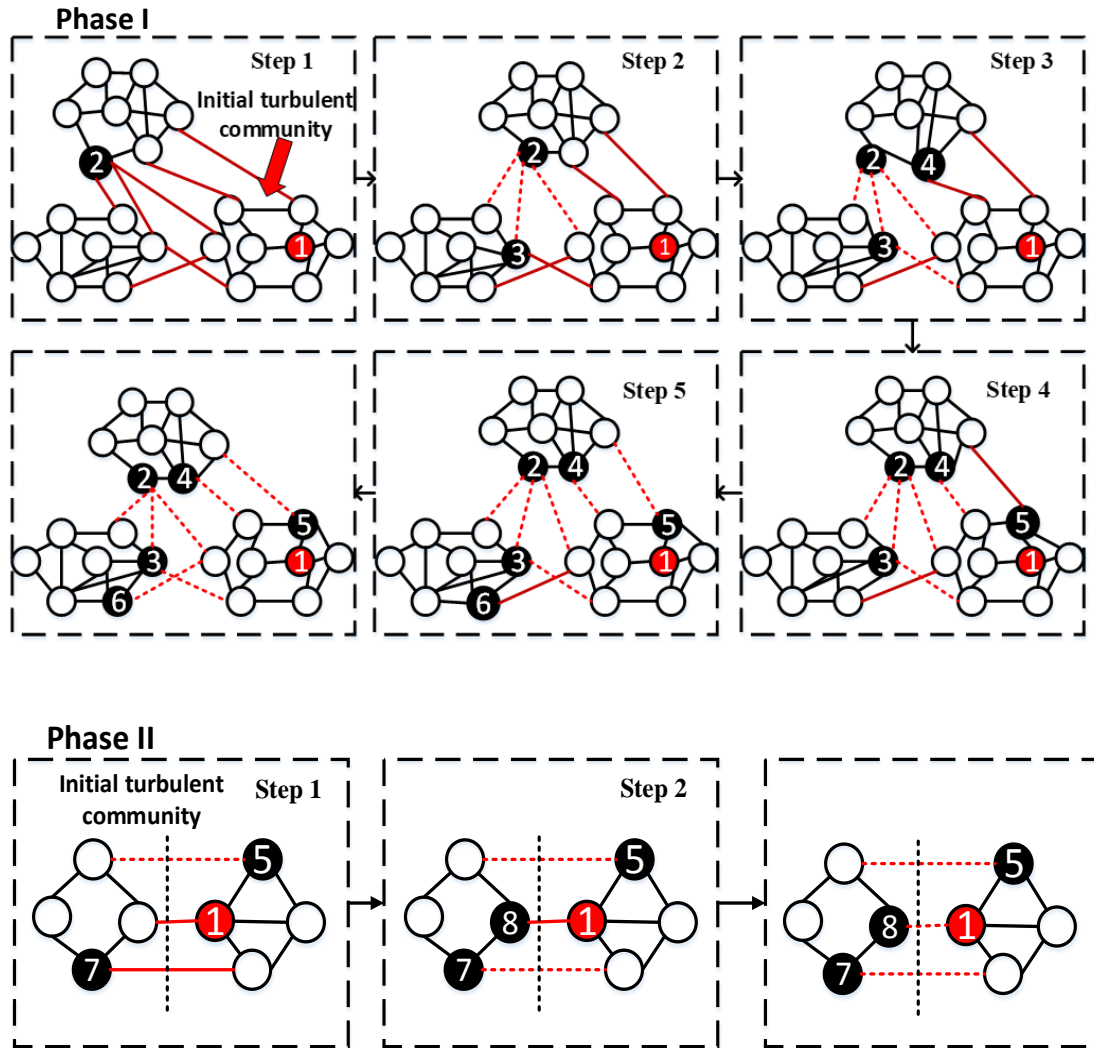


图 4-12 免疫过程示意图。Phase I 表示免疫第一阶段，Phase II 表示免疫第二阶段。免疫步长 $X = 1$ 。注：在每个 step 中计算业务度数时，每个子图中的虚线不再考虑。

仿真环境设置如下：考虑 BA 网络、WS 网络和随机网络（RD: Random Network）。网络节点数为 50，平均节点度大致保持相等。链路最大容量为 10Gbps。300 条业务流通过 Dijkstra 最短路径算法随机分布在网络中，即随机选择起始节点和目的节点。业务带宽需求遵循正态分布 $D \sim N(3, 0.1)$ ，随机分配业务优先级 $\varphi=1, 2, 3, 4, 5$ 。在实际网络中，一条业务可能是成百条小带宽业务的聚合，因此实验过程中设置的业务带宽需求和链路总容量可相比（同一数量级），且更容易产生网络拥塞，观测到的突变业务流的影响传播过程和免疫控制策略效果更为明显。不失一般性，随机从 300 条业务中选择突变业务流，并且将其带宽突增 5 倍。

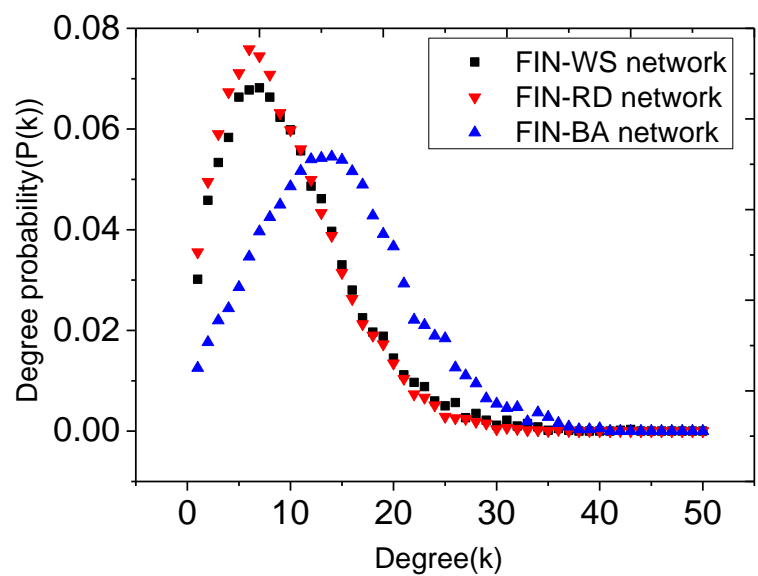


图 4-13 不同网络结构下业务流耦合关联网络的度分布

表 4-2 一级社团的结构属性

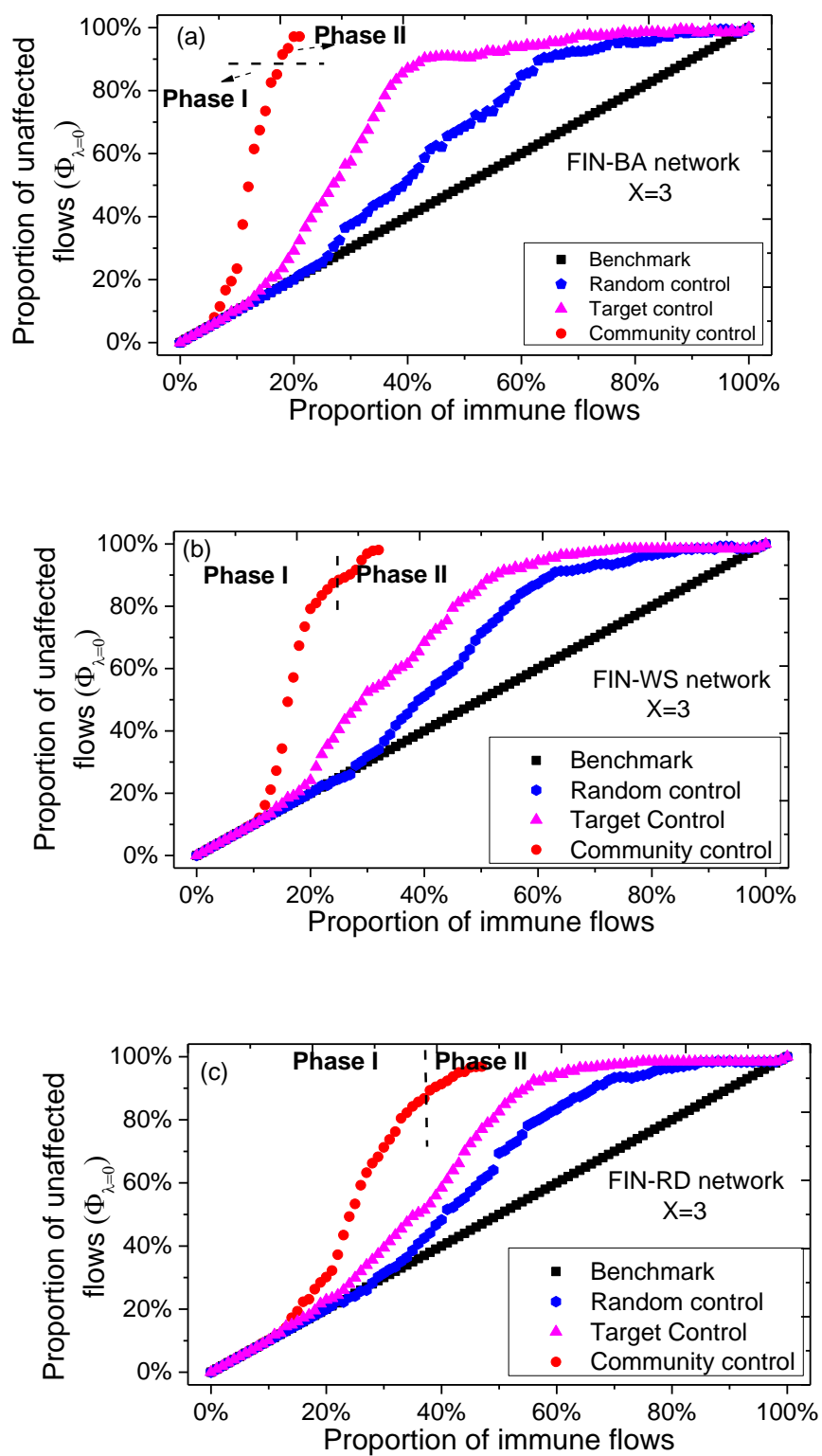
	FIN-BA	FIN-WS	FIN-RD
Flow number	300		
Community boundary flows	23%	34%	50%
Q	0.659	0.532	0.446

为简单起见，将从 BA 网络，WS 网络和 RD 随机网络得到的业务流耦合关联网络分别命名为 FIN-BA 网络（FIN: Flow Interaction network），FIN-WS 网络和 FIN-RD 网络。图 4-13 考察了不同 FIN 网络的度分布情况。在 FIN-WS 和 FIN-RD 网络中，业务度数大多集中在 20 以内。而在 FIN-BA 网络中，度分布范围较广，最高业务度数达到近 40。另外，表 4-2 显示了不同业务流耦合关联网络一级社团划分结果。从模块度 Q 值大小可以看出，FIN-BA 网络社团结构最为明显，模块度达到 0.6 以上，这与无标度网络高异质性有紧密关系。不同网络类型中，一级社团边界业务（Community boundary flows）的数目也不同，FIN-RD 网络最多，占到总业务数的 50%；FIN-BA 网络最少，只占总业务数的 23%。一般地，社团结构越明显的网络，其社团边界业务数目越少。表 4-2 可辅助分析后文仿真

得到的实验结果。

仿真考虑另外两种免疫控制策略作为比较：随机免疫策略（Random Immunization）和目标免疫策略（Target Immunization）。在随机免疫策略中，随机选择业务进行免疫。在目标免疫策略中，具有更高度数的业务优先被免疫。图 4-14 显示了不同免疫策略的传播控制效果。 x 轴代表免疫业务占总业务数的比例，在每个 Step 中免疫步长 $X = 3$ 。 y 轴代表免疫控制之后网络中带宽未受影响的业务比例 $\Phi_{\bar{\lambda}=0}$ 。图中“Benchmark”表示免疫业务数占总业务数的比例，可作为衡量免疫效果的基准线。不同免疫策略下 $\Phi_{\bar{\lambda}=0}$ 的仿真结果与基准线的偏差值大小可以反映出受保护业务的比例，夹角越大，控制免疫效果越好。在图 4-14 中可以看到，目标免疫控制策略比随机免疫控制策略表现更好。在目标免疫控制策略中，度数大的业务优先被免疫，度数大的业务其向外扩散的影响力相对也大，比随机地免疫业务更有效。此外，以往研究表明，目标免疫控制策略更适合于异构网络^[35-36]。图 4-13 所示的度分布中，FIN-BA 网络具有很多度数较大的业务，度分布最不均匀，比其他网络更为异构。因此，图 4-14 所示的目标免疫控制策略在 FIN-BA 网络中性能最好，在 FIN-RD 网络中性能最差。

基于社团的免疫控制策略在免疫第一阶段（图 4-14 中 Phase I）初期的控制效果并未明显优于其他控制策略，这是由于少量的免疫业务很难完全阻隔社团间的影响传播，难以形成以社团为单位的封闭保护区。但随着社团边界业务免疫比例的增长，控制效果迅速增加，并且明显优于其他策略。在 FIN-BA 网络中，18% 的社团边界免疫业务可以保护 85% 的业务不受突变业务流的影响。而为达到相同的结果，利用目标免疫控制策略需要 40% 的免疫业务，随机免疫控制策略则需要 60% 的免疫业务。在 FIN-WS 和 FIN-RD 网络中，对比数据变为 24%-51%-60% 和 37%-52%-63%。此外，基于社团的免疫控制策略在 FIN-BA 网络中效果最好。表 4-2 所示的业务流耦合关联网络社团属性中，FIN-BA 网络社团结构最为明显，社团边界业务最少，这使得免疫过程中分离出突变业务更容易。在免疫第二阶段（图 4-14 中 Phase II）对二级社团边界业务进行免疫以进一步缩小受影响的范围。第二阶段的免疫过程相较于第一阶段有很大程度的缩减。结合两个阶段共同免疫控制的结果，在 FIN-BA 网络中，21% 的社团边界免疫业务可以保护 97.2% 的业务免受带宽波动影响。在 FIN-WS 网络中，32% 的社团边界免疫业务可以保护 98.0% 的业务免受带宽波动影响。在 FIN-RD 网络中，47% 的社团边界免疫业务可以保护 96.8% 的业务免受带宽波动影响。为达到同等的控制效果，

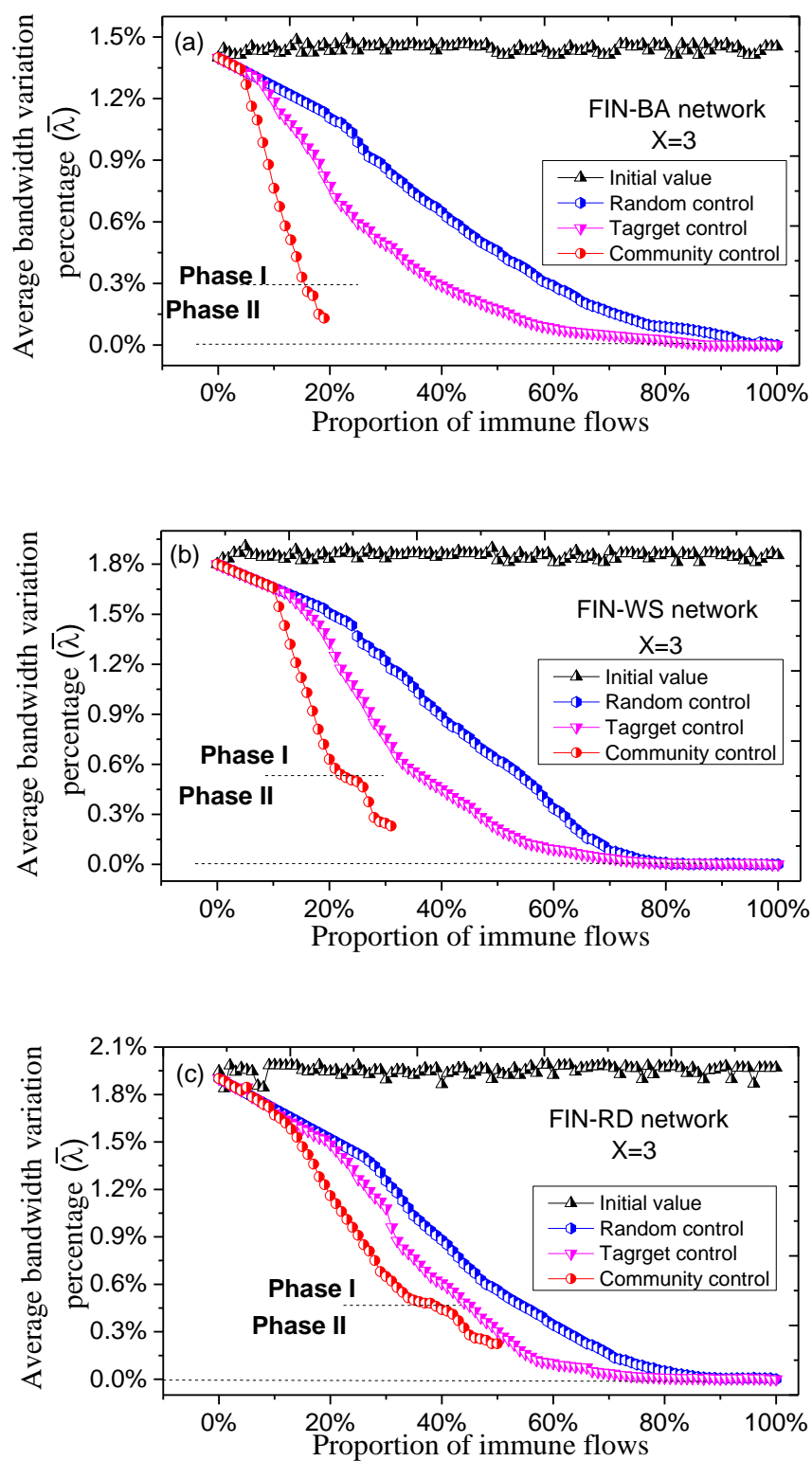
图 4-14 不同免疫策略下带宽未受影响的业务比例 $\Phi_{\lambda=0}$

随机免疫和目标免疫策略则需免疫 70%-90% 的业务。

从影响传播路径特点分析上述现象产生的原因。由于目标免疫过程中优先免疫影响力更大的中枢节点，从而保护范围更广，因此目标免疫比随机免疫效果更好。但网络中一旦存在大量的环路，影响便很容易绕过中枢节点而继续扩散。事实上，由于共享相同链路的业务是全连接的，FIN 网络包含大量的环路，因此目标免疫策略的控制效果失去优势。在 FIN 网络中控制带宽波动传播更有效的方法是形成一个封闭的免疫区域，并阻断所有可能的传播路径，这就是为什么将业务划分为社团结构并免疫社团边界业务的原因。表 4-2 中，一级社团边界业务的占比约为 23%~50%，这意味着可以通过免疫不超过 23%~50% 的业务来粗略地隔离一级社团间的影响传播。理论上免疫业务的数量可以继续减少一半（即 11.5%~25%），这是由于社团之间的连边只需有一个端点被免疫便可以阻断传播。在本文的免疫策略制定中已考虑到这一点，因此基于社团结构的免疫策略所需免疫业务数目远远低于其他策略，相比于比随机免疫控制策略和目标免疫控制策略 具有明显的优势。

由于任何免疫控制方案都不能保证网络所有业务均能受到保护，因此，继续利用业务突变前后全网业务带宽平均受影响程度 $\bar{\lambda}$ 来衡量不同策略的控制效果。如图 4-15 所示，“Initial Value”表示未经过免疫控制时全网业务平均带宽变化率，即初始带宽变化率。FIN-BA 网络中平均业务度数最大（图 4-13），依据公式（4-11）至公式（4-15），度数越大表明越多的业务将同时分担带宽变化，可有效降低突变业务流的影响强度。所以无论在哪种策略下，FIN-BA 网络中的平均业务受影响程度均小于 FIN-WS 和 FIN-RD 网络。针对同一 FIN 网络，基于社团的免疫控制策略在减小带宽变化方面比其他策略表现更好。由于基于社团的免疫控制策略可以保护更多的业务免受影响，因此在相同数量的免疫业务下，网络业务平均带宽变化率最小。具体数据显示，与初始带宽变化率相比，经过 phase I，业务平均带宽变化率降至 0.3%~0.5%，经过 phase II，业务平均带宽变化率进一步降低至 0.13%~0.15%。由于本文采取的策略仅免疫了一级社团和二级社团的边界业务，因此未能像其他免疫策略一样大面积，但在相同数量的免疫业务下，基于社团的免疫控制策略使得网络业务平均带宽变化率最小。

最后，本文讨论了社团免疫控制策略中免疫步长 X 与控制效率之间的关系。免疫步长 X 表示在每个 step 中，免疫社团边界业务的数目。以 FIN-BA 网络为例，图 4-16 显示了免疫第一阶段（phase I）中不同免疫步长策略下的控制效果。

图 4-15 不同免疫策略下业务带宽平均受影响程度 $\bar{\lambda}$

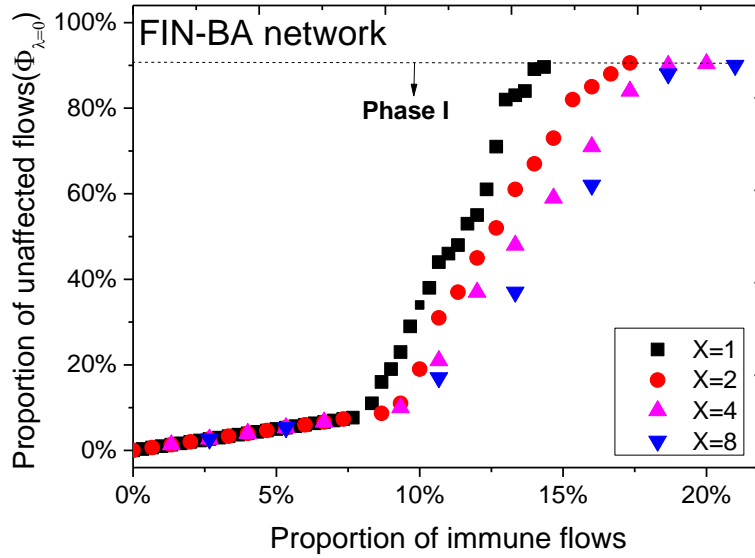


图 4-16 免疫步长与控制效率之间的关系

可以看出，在相同免疫业务数目下，随着 X 的增长，网络中未受影响的业务比例 ($\Phi_{\lambda=0}$) 逐渐变小，控制效率降低。上文指出，如果社团之间连边的一端被免疫，影响便不能通过该连接进行传播。但如果在同一个免疫步骤中越多的业务被免疫，这将增加社团间连边两个端点同时被免疫的可能性，从而降低了免疫策略的效率。在图 4-16 中，当免疫步长 $X=1$ 时，14.3% 的免疫业务便可以隔离一级社团。而为达到相同的隔离效果，当免疫步长为 $X=2, 4, 8$ 时，则需要 15%~20% 的免疫流。因此，可以得出如下结论：免疫步长越小，算法复杂度提升，但控制效果更显著。

4.3 本章小结

本章首先建立了突变业务流的全网影响传播模型。为清晰反映传播过程，构建了具有层次结构的业务流耦合关联网络，使用“邻居阶数”描述突变业务流和其余业务间的邻近关系，并且定义“耦合概率”描述相邻两阶邻居业务的耦合作用。网络拥塞环境下，突变业务流的传播行为特征总结如下：1) 突变业务流对其奇数阶邻居有负面影响，会使奇数阶邻居业务带宽减少；对偶数阶邻居有正面影响，会使偶数阶邻居业务带宽增加；2) 突变业务流的影响强度随着邻居阶数的增加而急剧下降，实验中相邻邻居阶数间的下降幅度可达数十倍以上；3) 随

着网络拥塞程度的增加, 突变业务流的影响强度也显著增加, 尤其是对一阶邻居的影响。本文进一步考虑网络结构变化带来的影响: 1) 当网络中业务数目固定时, 网络规模的扩大可以降低突变业务流对全网业务的平均影响强度; 2) 在相同网络规模和随机业务分布情况下, 由于小世界网络中业务之间的耦合概率较低, 因此, 在降低突变业务流的影响强度方面优于随机网络和无标度网络。

考虑到业务流耦合关联网络的社团属性, 本章提出了一种基于社团结构的免疫控制策略, 有效控制突变业务流造成的全网业务带宽动荡现象。仿真结果表明, 由于业务流耦合关联网络中存在大量环路, 因此基于社团的免疫控制策略优于随机免疫策略和目标免疫策略。以无标度网络中的业务为例, 21%的免疫业务可以保护 97.2%的业务免受带宽波动影响。为达到同样的结果, 其他控制策略则需要 70%-90%的免疫业务。此外, 通过合理调节在每个控制步骤中免疫业务的数目, 可以显著改善控制效果, 但相应的算法复杂度也会提高。

参考文献

- [1] Holme, P. Temporal network structures controlling disease spreading[J]. *Physical Review E*, 2016, 94(2-1):022305.
- [2] He, Z., Cai, Z., Yu, J., et al. Cost-efficient Strategies for Restraining Rumor Spreading in Mobile Social Networks[J]. *IEEE Transactions on Vehicular Technology*, 2017, 99:1-1.
- [3] Biedermann, S., Katzenbeisser, S. Detecting Computer Worms in the Cloud[J]. *Lecture Notes in Computer Science*, 2017, 7039:43-54.
- [4] Wang, J., Jiang, C., Qian, J. Robustness of Internet under targeted attack: A cascading failure perspective[M]. Academic Press Ltd. 2014.
- [5] Najjari, N., Min, G., Hu, J., et al. Performance Analysis of WLANs with Heterogeneous and Bursty Multimedia Traffic[C]// *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*. IEEE, 2018:1-6.
- [6] Miao, W., Min, G., Wu, Y., et al. Performance Modelling and Analysis of Software-Defined Networking under Bursty Multimedia Traffic[J]. *Acm Transactions on Multimedia Computing Communications & Applications*, 2016, 12(5):77-97.

- [7] Jiang, H., Dovrolis, C. Source-level IP packet bursts: causes and effects[J]. Proc Imc, 2003:301-306.
- [8] Nguyen, H., Liu, W., Chen, F. Discovering Congestion Propagation Patterns in Spatio-Temporal Traffic Data[J]. IEEE Transactions on Big Data, 2017, 99:1-1.
- [9] Ohsaki, H., Sugiyama, K., Imase, M. Congestion Propagation among Routers with TCP Flows[J]. International Journal of Computer Networks & Communications, 2009, 1(2):112-128.
- [10] St éger, J., Vadera, P., Vattay, G. On the propagation of congestion waves in the Internet[J]. Physica A Statistical Mechanics & Its Applications, 2006, 360(1):134-142.
- [11] 王成, 刘金刚, 刘汉武. 网络中突发业务自相似建模及其 Hurst 系数估计[J]. 计算机工程, 2006, 32(2):101-103.
- [12] Ducrot, A., Langlais, M., Magal, P. Qualitative analysis and traveling wave solutions for the SI model with vertical transmission[J]. Communications on Pure & Applied Analysis, 2017, 11(1):97-113.
- [13] Ge, J., Lin, Z., Zhu, H. Environmental risks in a diffusive SIS model incorporating use efficiency of the medical resource[J]. Discrete and Continuous Dynamical Systems - Series B (DCDS-B), 2017, 21(5):1469-1481.
- [14] Khoshnaw, S., Mohammad, N., Salih, R. Identifying Critical Parameters in SIR Model for Spread of Disease[J]. Open Journal of Modelling & Simulation, 2017, 5(1):32-46.
- [15] Cai, Y., Kang, Y., Wang, W. A stochastic SIRS epidemic model with nonlinear incidence rate[J]. Applied Mathematics & Computation, 2017, 305:221-240.
- [16] Fekete, A., Vattay, G., Kocarev, L. Traffic dynamics in scale-free networks[J]. Complexus, 2006, 3(1-3):97-107.
- [17] Currently, S. A Case Study of IPv6 Network Performance: Packet Delay, Loss, and Reordering[J]. Mathematical Problems in Engineering, 2017, 2017(4):1-10.
- [18] Han, Y., Moutarde, F. Analysis of Large-Scale Traffic Dynamics in an Urban

- Transportation Network Using Non-Negative Tensor Factorization[J]. International Journal of Intelligent Transportation Systems Research, 2016, 14(1):36-49.
- [19] Park, H. W., Chung, J. W. A study on reduction of traffic burstiness using window based segment spacing[A].//Proceedings of International Conference on Information Networking [C]. Washington: IEEE, 2001:41-45.
- [20] Torabzadeh, M., Ji, Y. S., A Traffic Engineering-QoS Approach to the Traffic Burstiness in Short-Time Scales[A].//International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06) [C]. Moren: IEEE, 2006: 23-29.
- [21] Sathiaselvan, A., Fairhurst, G. TCP-Friendly Rate Control (TFRC) for bursty media flows[M]. Elsevier Science Publishers B. V. 2011.
- [22] 张鹤颖, 窦文华. 主动队列管理机制在突发流下的性能比较与分析[J]. 计算机工程与科学, 2004, 26(12):8-11.
- [23] Zhang, X., Wu, J., Zhao, P., et al. Epidemic spreading on a complex network with partial immunization[J]. Soft Computing, 2017:1-9.
- [24] Buono, C., Braunstein, L. A. Immunization strategy for epidemic spreading on multilayer networks[J]. EPL, 2015, 109(2):26001.
- [25] Zuzek, L. G. A., Buono, C., Braunstein, L. A. Epidemic spreading and immunization strategy in multiplex networks[J]. Journal of Physics Conference Series, 2015:49-56.
- [26] Wu, X., Liu, Z. How community structure influences epidemic spread in social networks[J]. Physica A Statistical Mechanics & Its Applications, 2008, 387(2):623-630.
- [27] Salathé M., Jones, J. H. Dynamics and control of diseases in networks with community structure[J]. 2010, 6(4):387-395.
- [28] Shang, J., Liu, L., Li, X., et al. Epidemic spreading on complex networks with overlapping and non-overlapping community structure[J]. Physica A Statistical

- Mechanics & Its Applications, 2015, 419(419):171-182.
- [29] Gupta, N., Singh, A., Cherifi, H. Community-based immunization strategies for epidemic control[A].// International Conference on Communication Systems and Networks[C]. Bangalore: IEEE, 2014:1-6.
- [30] Stegehuis, C., Hofstad, R. V. D., Leeuwaarden, J. S. H. V. Epidemic spreading on complex networks with community structures[J]. Scientific Reports, 2016, 6:29748.
- [31] Liu, T., Li, P., Chen, Y., et al. Community Size Effects on Epidemic Spreading in Multiplex Social Networks[J]. Plos One, 2016, 11(3):e0152021.
- [32] Newman, M. E. J., Girvan, M. Finding and evaluating community structure in networks[J]. Phys Rev E Stat Nonlin Soft Matter Phys, 2004, 69(2):026113.
- [33] Newman, M. E. J. Fast algorithm for detecting community structure in networks[J]. Phys Rev E Stat Nonlin Soft Matter Phys, 2003, 69(6 Pt 2):066133.
- [34] Blondel, V. D., Guillaume, J. L., Lambiotte, R., et al. Fast unfolding of community hierarchies in large networks[J]. J Stat Mech, 2008, abs/0803.0476.
- [35] Tanaka, G., Urabe, C., Aihara, K. Random and Targeted Interventions for Epidemic Control in Metapopulation Models[J]. Scientific Report, 2014, 4(4):05522.
- [36] Xia, L., Jiang, G., Song, Y., et al. An improved local immunization strategy for scale-free networks with a high degree of clustering[J]. European Physical Journal B, 2017, 90(1):2.

第五章 基于流过载失效的 网络故障传播模型与业务流动态行为研究

网络中的元素通常高度相互依赖，一次小范围的攻击有时便可破坏整个系统，因此，针对网络级联故障的研究备受关注^[1-10]，如电力网络的大面积停电事故^[11-12]、交通网络高峰期的异常拥堵^[13-14]等。同样地，通信网络也极易受到网络故障的影响，业务突发或路由器的损坏，均可能会引发网络级联失效^[15-16]。了解业务在网络故障发生时的行为特性十分必要，直接影响网络服务质量的保障和提升^[17-18]。由于网间流量巨大，且与网络元件（路由器和链路）之间的耦合关系复杂，级联故障中业务行为的建模与分析极为困难。

一些学者针对基于业务过载的网络级联故障进行了研究。Crucitti 等人^[19]提出了基于业务流动态再分配的级联失效模型，研究发现当某一节点的邻居节点均为超大负荷节点时，该节点的故障便足够引起整个系统的崩溃。Guo 等人^[20]基于复杂系统的“临界自组织性”（Self-organized Criticality）和“高容忍性”（Highly Optimized Tolerance）提出了基于网络动态协议和流量负载的故障模型，分析故障期间网络结构的 RYF 相变（Robustness yet Fragile）。RYF 经常用来描述网络在健壮性和脆弱性间转变的性质^[21-22]。仿真结果表明，互联网的 RYF 相变与异常的网络负载模式密切相关。Liu 等人^[23]发现单节点（路由器）故障会增加其邻居节点的流量负载压力，导致路由器网络级联失效。不难发现，目前已有研究多集中于故障过程中网络宏观行为的研究，例如网络效率和网络拓扑相变等问题，尚未对级联故障中业务流的行为进行系统地建模与分析。

本章将针对业务流过载导致的链路失效问题建立网络级联故障模型，并重点研究故障传播过程中业务流的动态行为特征。具体而言，基于业务流耦合关联网络，增加其与实际物理网络的关联关系，构建业务层/网络层双层网络模型。基于该双层网络，一方面研究故障中业务传输与路由器、链路之间的支撑依赖关系变化；另一方面分析级联故障中业务本身的动态行为变化，如业务传输行为、拥塞竞争行为等，反映网络级联失效对网络业务服务质量的影响。

5.1 渗流理论基本原理

渗流理论（Percolation theory）是研究复杂网络鲁棒性的基本方法。将网络

中的部分节点和节点相连接的链路依次去除，这一过程就称为“渗流”，或者更准确的称为“点渗流”。渗流可以作为现实世界中许多现象的模型，例如 Internet 中的路由器故障就可以看作相应的节点和链路从网络中去除的过程。在这一过程中，路由器故障会导致该路由器无法接收数据，同时，经过该路由器的数据也无法到达其他路由器，使网络通信不得不选择其他更长或更拥塞的路径，甚至造成部分区域的断网。渗流理论可帮助理解节点或链路故障导致的网络级联效应。

渗流理论的基本问题描述如下：假设网络中有 $1-p$ 比例的节点失效（人为攻击或自然灾害），与失效节点相连的链路也随之失效，进而网络被分割成若干连通片以及孤立的节点。通常假设只有最大连通片中的节点具有功能，最大连通片所占比例 P_∞ 则是衡量网络鲁棒性的指标。

Newman, Strogatz 和 Watts 首次提出了随机网络中基于网络发生函数（Generating Function）的渗流理论^[24]，奠定了复杂网络渗流理论研究的基础。

给定网络 i 的规模为 N ， p_k 为节点度为 k 的概率，则网络度分布函数定义为：

$$G_i(x) = \sum_{k=0}^{\infty} p_k x^k \quad (5-1)$$

其中， x 为随机变量。网络的平均度为：

$$\langle k \rangle = G'_i(1) \quad (5-2)$$

在渗流理论中较为重要的一个参量是网络连通子图的大小，与连通子图相关的网络函数为余度分布函数，或称作分支函数。分支函数 $H_i(x)$ 定义为：

$$H_i(x) = \frac{G'_i(x)}{G_i(1)} = \frac{\sum_{k=1}^{\infty} k p_k x^{k-1}}{\langle k \rangle} \quad (5-3)$$

分支函数可以理解为每个节点以概率 $\frac{k p_k}{\langle k \rangle}$ 和一个度为 k 的节点连边的过程。

下面介绍如何求解网络最大连通片的大小。对于随机删除 $1-p_i$ 比例节点的网络 i ，最大连通片中节点占网络总节点的比例满足：

$$P_{\infty,i} = p_i g_i(p_i) \quad (5-4)$$

其中, p_i 满足:

$$g_i(p_i) = 1 - G_i(p_i f_i(p_i) + 1 - p_i) \quad (5-5)$$

$$f(p_i) = H_i(p_i f_i(p_i) + 1 - p_i) \quad (5-6)$$

对于恶意删除 $1 - p_i$ 比例节点的网络 i , $P_{\infty,i}$ 计算公式如下:

$$P_{\infty,i} = 1 - G'_i(1 - p_i(1 - f_i)) \quad (5-7)$$

其中, $G'_i(x)$ 为剩余节点得到的等价网络生成函数, 满足下式:

$$P_{\infty,i} = 1 - G'_i(1 - p_i(1 - f_i)) \quad (5-8)$$

其中, $G'_i(x)$ 为等价网络的分支函数。

P_{∞} 随 p 的函数变化可体现出最大连通片的形成和解体的过程, 称之为“渗流相变”。目前, 相变类型主要分为一级不连续相变和二级连续相变。在不连续相变过程中, 会出现网络瞬间瓦解的情况, 而在连续相变中, 网络逐渐解体。 P_{∞} 由零值向非零值转变时的 p 值, 称作“渗流阈值” p_c 。 p_c 值越小, 网络鲁棒性越高。

5.2 基于流过载的双层网络级联故障模型

5.2.1 业务层/网络层双层网络模型

网络耦合行为不仅体现在业务服务质量之间, 也体现在业务与物理链路、节点之间。节点和链路依据通信协议对业务进行约束和控制, 业务负载分布反之会影响节点和链路的性能, 这种耦合关联在网络发生级联故障时表现的尤为明显: 当网络受到不可抗力影响时, 如设备老化、自然灾害、网络恶意攻击, 节点和链

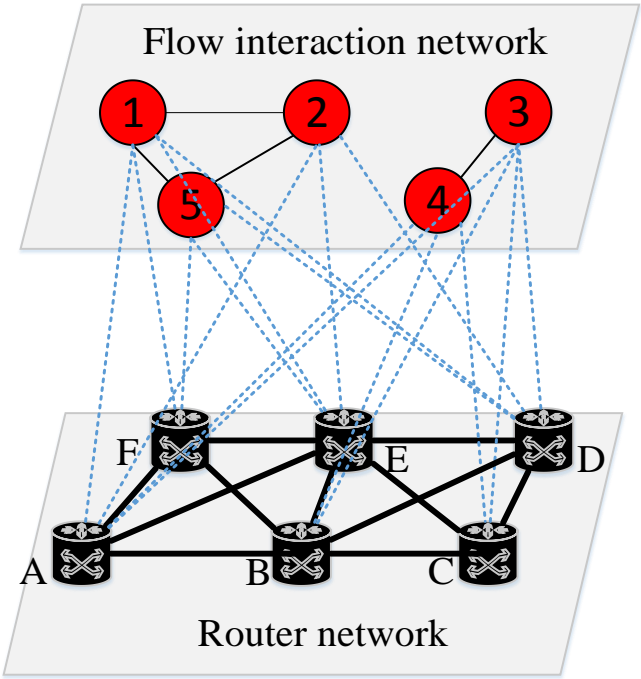


图 5-1 业务层/网络层双层网络

表 5-1 业务路径表

Service	Path
1	A→F→E→D
2	A→E→D
3	A→B→C→D
4	A→B→C
5	F→E→D

路之间的连接关系将发生改变，网络拓扑的变化将导致与受损节点或链路相关联的业务传送中断，业务按照剩余节点组成的网络重新选路。网络动态变化会导致网络资源重新分配，业务负载发生转移。而节点或链路的过载使得设备在网络拥塞缓解之前，不能继续承载多余业务，进而产生暂时性的功能失效。这将导致新一轮的网络资源重新分配与设备过载失效，引发网络级联故障。可见，网络对业务传输具有支撑作用，而业务负载分配也影响网络性能的高低，两者之间存在相互依赖和影响的关系。

为清晰反映业务传输与承载网络间的关系，本节建立了业务层/网络层双层

网络模型，其中业务层由业务流耦合关联网络表示，网络层由路由器网络表示，业务层和网络层的跨层连接表示业务与路由器间的支撑依赖关系。举例说明，图 5-1 显示了路由器网络(Router Network)与业务流耦合关联网络(Flow Interaction Network)共同构建的双层网络，业务流耦合关联网络依据业务路径表 5-1 建立，存在共享链路的两个业务之间有连边。跨层连接满足以下规则：当业务通过某路由集合时，将该业务和路由集合中的路由器建立无向连边。例如业务 3 经过路由器 $A \rightarrow B \rightarrow C \rightarrow D$ (见表 5-1)，因此，将业务 3 与路由器 A 至 D 分别建立连接(见图 5-1 中虚线)。

对于业务流耦合关联网络 $G_f = (V_f, E_f)$ ，节点集 V_f 表示网络中全体业务流集合，边集 E_f 表示业务间竞争关系。节点权重矩阵 $W_f = \{w_f\}$ 量化业务拥塞程度：

$$w_f = \frac{d_f - b_f}{d_f} \quad (5-9)$$

其中， d_f 为业务发送速率或带宽需求， b_f 为业务实际传送速率或传输带宽。

对于路由器网络 $G_r = (V_r, E_r)$ ， V_r 代表路由器集合， E_r 代表链路集合。 $W_r = \{w_r\}$ 是量化链路传输容量的边权矩阵：

$$w_r = l \quad (5-10)$$

其中， l 为链路传输容量。

5.2.2 流过载级联故障模型

流量过载引发的网络级联失效是常见的网络故障现象之一，例如交通系统中，地铁网络某些节点的故障会导致整条线路其他节点的停运或人流量拥挤，进而造成更大规模的停运。基于流过载的网络级联失效通常表示为如下数学过程：考虑具有 N 个节点的加权无向无网络 G 。节点 i 的负载 L_i 由介数中心性决定。节点容量定义为^[25]：

$$C_i = (1 + \alpha)L_i \quad (5-11)$$

参数 $\alpha \in [0, 1]$ 为节点容量的调节系数。如果节点由于故障或攻击从网络中移除后，移除节点会影响其他节点之间的最短路径，网络负载将重新分布，若此时某些节点的负载大于其容量，那么节点将失效并从网络中移除。这一过程周而复始，直至网络中所有节点的负载均小于其容量限制。

以往研究中网络级联故障通常由节点连锁故障引起：当负载超过节点的最大承受能力时，节点就会产生故障，进而导致业务重分配，使得更多的节点失效，最终引起网络级联故障。节点故障模型并不完全符合通信网络中路由器实际工作原理。路由器多个端口并行工作，链路容量的大小决定了每个端口的最大转发速率。如果数据包超过了每个端口单位时间的处理能力，那么数据包将在缓冲区中等待，从而导致链路拥塞。当拥塞达到一定程度时，或为严重负载时，端口便不能正常工作，导致相对应的链路失去正常传输功能。因此，在实际网络中，拥塞会造成相应链路故障，而并非作用在整个路由器上。在本节流过载失效问题中，设备失效是基于链路的失效，而非节点的失效。

本节提到的故障是指功能故障，当网络拥塞程度超过设备处理能力时，设备将无法承载后续任务要求，产生功能失效。通常，网络会通过各种调节机制来恢复设备，但由于恢复机制非常复杂，如缓冲机制和队列调度机制^[26]，需要一段时间才能产生效果，在这段时间内故障可能会继续蔓延传播，对全网造成不可恢复的灾难，本节则主要考虑在网络设备尚未恢复功能前的级联故障过程。

根据上述对链路过载失效原理的描述，现将网络级联故障过程总结如下：由于网络业务资源的不合理分配，导致某些链路过载超出最大处理能力，链路产生功能性故障（短时间不可恢复）。链路失效使得后续到达的业务须绕过故障链路选择备用路径传输，导致业务资源重新分配，进而可能致使其他链路的过载和失效。整个过程循环往复，网络发生级联故障。

业务重分配导致的链路过载是产生网络级联故障的直接因素，这里首先介绍业务负载的重分配过程。

在路由器网络 $G_r = (V_r, E_r)$ 中， $L = \{l\}$ 表示链路的传输容量。在业务流耦合关联网中 $G_f = (V_f, E_f)$ ， $D = \{d\}$ 表示业务带宽需求或发送速率。 $B = \{b\}$ 是实际的业务传输带宽或传送速率。那么 $d - b$ 则是单位时间在缓冲器中等待（拥塞）的数据包数量。显然，传输带宽不会超过带宽需求，即 $b \leq d$ 。 $\Gamma = \{\varphi\}$ 表示业务优先级。优先级高的业务可优先传输，业务传输带宽与其优先级成比例^[27-28]。

对于每个业务，一旦确定其源地址和目的地址，便可根据特定的路由算法选择出一条从源到目的地的路径。如果有多条路径满足条件，随机选择其中一条。假设网络中任一业务 f_o 的传输路径表示为其所经链路构成的集合 $\{e_{r,1}, e_{r,2}, \dots, e_{r,i}, \dots, e_{r,N}\}$ 。基于比例带宽分配算法^[27-28]，链路 $e_{r,i}$ 允许 f_o 传输的带宽为：

$$b_{f_o, e_{r,i}} = \min \left\{ l_{e_{r,i}} \frac{\varphi_{f_o}}{\sum_{f_o \in F_{e_{r,i}}} \varphi_{f_o}}, d_{f_o} \right\} \quad (5-12)$$

其中, $l_{e_{r,i}}$ 是链路 $e_{r,i}$ 的传输容量。 φ_{f_o} 和 d_{f_o} 分别表示业务 f_o 的优先级和带宽需求。

$F_{e_{r,i}}$ 是所有经过链路 $e_{r,i}$ 的业务流集合。

同时, 业务 f_o 的传输带宽也受到传输路径上其他链路容限的影响。基于 TCP/IP 协议的反馈机制, 业务的传输带宽在其路径上是连续的, 所以业务 f_o 的最终传输带宽由最小允许带宽决定:

$$b_{f_o} = \min \{ b_{f_o, e_{r,1}}, b_{f_o, e_{r,2}}, \dots, b_{f_o, e_{r,i}}, \dots, b_{f_o, e_{r,N}} \} \quad (5-13)$$

对于链路 $e_{r,i}$, 传输负载定义为:

$$l'_{e_{r,i}} = \sum_{f_o \in F_{e_{r,i}}} b_{f_o} \quad (5-14)$$

利用公式 (5-12) - (5-14), 可以得出所有业务的传输带宽以及链路的负载情况, 下面进一步研究链路的故障判定。

如果路由器端口到达的数据包超过了单位时间能够处理的数量, 那么额外的数据包会在缓冲区中等待并建立队列。根据 TCP / IP 协议^[29], 缓冲区大小由链路容量和分组往返时间 (Round-Trip Time, RTT) 决定。参数 β 反映数据包的平均往返时间, 缓冲区大小定义为:

$$Q_{e_{r,i}} = \beta l_{e_{r,i}} \quad (5-15)$$

如果链路 $e_{r,i}$ 中的数据拥塞量大于缓冲区 $Q_{e_{r,i}}$ 大小, 即满足下式时:

$$\sum_{f_o \in F_{e_{r,i}}} d_{f_o} - l'_{e_{r,i}} > Q_{e_{r,i}} \quad (5-16)$$

链路 $e_{r,i}$ 中业务数据产生溢出, 链路 $e_{r,i}$ 过载失效; 否则, 链路 $e_{r,i}$ 正常工作。依据

公式 (5-9), (5-14), (5-15), 上式可改写为:

$$\sum_{f_o \in F_{e_{r,i}}} d_{f_o} w_{f_o} > \beta w_{e_{r,i}} \quad (5-17)$$

其中， w_{f_o} 为业务 f_o 的节点权重， $w_{e_{r,i}}$ 为链路 $e_{r,i}$ 的边权重。

通过上述建立的链路失效规则，可确定当前网络发生故障的链路。将失效链路移除网络后，业务会重新分配，进而导致新一轮的级联失效过程。因此，总结网络级联失效过程如下：

Step1: 网络由于自身原因或外部原因，产生业务负载异常，导致一部分节点和链路失效；

Step2: 由于故障链路不能继续正常承担业务，暂时将它们从网络中删除（非永久性删除），更新路由器网络拓扑结构 $G_r = (V_r, E_r)$ 和边权矩阵 $W_r = \{w_r\}$ ；

注：链路失效可能会导致孤立节点的产生，同样进行删除；

Step3: 网络业务根据新的网络结构重新路由并找到备用路径，利用公式（5-12）的带宽分配算法，业务负载重新分配。依据新的业务关联关系，更新业务流耦合关联网络 $G_f = (V_f, V_f)$ 和点权矩阵 $W_f = \{w_f\}$ ；

Step4: 利用判定公式（5-17），判定新的失效链路。

Step5: 重复 Step2-4，直至网络负载分配不再导致新的链路产生故障，网络级联故障效应停止。

5.2.3 故障中业务动态行为模型

基于建立的流过载级联故障模型，进一步分析级联故障中业务的动态行为。

在业务层/网络层的双层网络中（图 5-1），业务和路由器之间的跨层连接可反映当前网络业务的传输效率。定义业务对网络的依赖强度 ρ ：

$$\rho(t) = \frac{\mu(t)}{N_R N_F} \quad (5-18)$$

t 表示故障过程中业务负载重新分配的次数，可以反映故障失效的时间进程。 N_R 表示当前网络中路由器的数量， N_F 表示当前网络中传输业务的数量。 $N_R N_F$ 表示所有可能的跨层连接数量， $\mu(t)$ 是实际存在的跨层连接数量，在网络级联失效过程中动态变化。对每条业务而言，传输路径上的路由器数目等于路径长度值加

1, 从而 $\mu(t)$ 表示为:

$$\mu(t) = \sum_{f_i \in N_F} (\text{len}(f_i, t) + 1) \quad (5-19)$$

其中, $\text{len}(f_i, t)$ 表示业务 f_i 的路径长度, 是随时间变化的变量。因此, ρ 可以表示为:

$$\rho(t) = \frac{\sum_{f_i \in N_F} (\text{len}(f_i, t) + 1)}{N_R N_F} \quad (5-20)$$

当业务数目足够大时, 用平均路径长度代替实际路径长度, 进一步得到 ρ 的理论值 $\rho^*(t)$:

$$\rho^*(t) = \frac{(\overline{s(t)} + 1)N_F}{N_R N_F} = \frac{\overline{s(t)} + 1}{N_R} \propto \overline{s(t)} \quad (5-21)$$

$\overline{s(t)}$ 是业务的平均路径长度。通常, ρ^* 值越低, 表示业务具有较短的平均路径长度, 传输效率更高。

除业务的传输效率外, 对网络故障中业务间的竞争行为也进行建模。网络最大连通片 g 通常被用来衡量级联故障对网络结构的影响^[30-34]。在路由器网络中, g 值反映路由器网络结构鲁棒性的变化, 表示为:

$$g_R = \frac{N'_R}{N_R} \quad (5-22)$$

其中, N'_R 是路由器网络中处在最大连通片中的节点数目。

同样地, 在业务流耦合关联网络中, 最大连通片 g_F 可以体现业务的最大的竞争区域。依据第三、四章对业务间竞争耦合关系的建模和分析, 最大连通片 g_F 内的业务在网络动态变化时相互关联、相互竞争。 g_F 公式表示为:

$$g_F = \frac{N'_F}{N_F} \quad (5-23)$$

其中, N'_F 是业务流耦合关联网络中处在最大连通片中的业务数目。

此外，定义业务竞争强度 γ ：

$$\gamma(t) = \frac{1}{N_F} \sum_{i=1}^{N_F} \frac{k_i(t)}{k_{\max}} w_{f_i}(t) \quad (5-24)$$

其中， $k_i(t)$ 表示业务 f_i 的连接度，随时间动态变化； k_{\max} 表示网络中业务的最大度数。归一化数值 $k_i(t)/k_{\max}$ 量化业务 f_i 与其邻居间的竞争紧密程度。 w_{f_i} 表示业务 f_i 的拥塞程度，业务越拥塞对带宽资源的竞争更强烈，因此， w_{f_i} 可以衡量业务 f_i 的竞争能力。根据 w_{f_i} 定义， γ 推导为：

$$\gamma(t) = \frac{1}{k_{\max} N_F} \sum_{i=1}^{N_F} k_i(t) \left(1 - \frac{b_{f_i}(t)}{d_{f_i}}\right) \quad (5-25)$$

γ 的最大值 $\gamma_{\max}(t)$ 为：

$$\begin{aligned} \gamma_{\max}(t) &= \frac{1}{k_{\max} N_F} \left(1 - \frac{b_{\min}(t)}{d_{\max}}\right) \sum_{i=1}^{N_F} k_i(t) \\ &= \frac{\overline{k(t)}}{k_{\max}} \left(1 - \frac{b_{\min}(t)}{d_{\max}}\right) \end{aligned} \quad (5-26)$$

$\overline{k(t)}$ 表示业务平均度， $b_{\min}(t)$ 表示当前网络中业务传输的最小传输带宽， d_{\max} 表示网络业务中的最大带宽需求。

以上，对网络级联失效过程中业务传输行为、竞争行为进行建模。下节将利用业务行为模型，对不同网络类型、故障类型下的业务动态行为进行仿真分析。

5.3 实验仿真与结果分析

5.3.1 网络环境与参数设置

网络设置：考虑两种路由器网络 BA 网络和 ER 网络。对于 BA 无标度网络，网络规模 $N_{BA} = 2000$ ， $\lambda = 2.6$ 。对于 ER 随机网络， $N_{ER} = 2000$ ，平均度为 $\bar{k}_{ER} = 10$ 。定义链路容量 l 为可变量， α 为调节系数：

$$l = \alpha l_0 \quad (5-27)$$

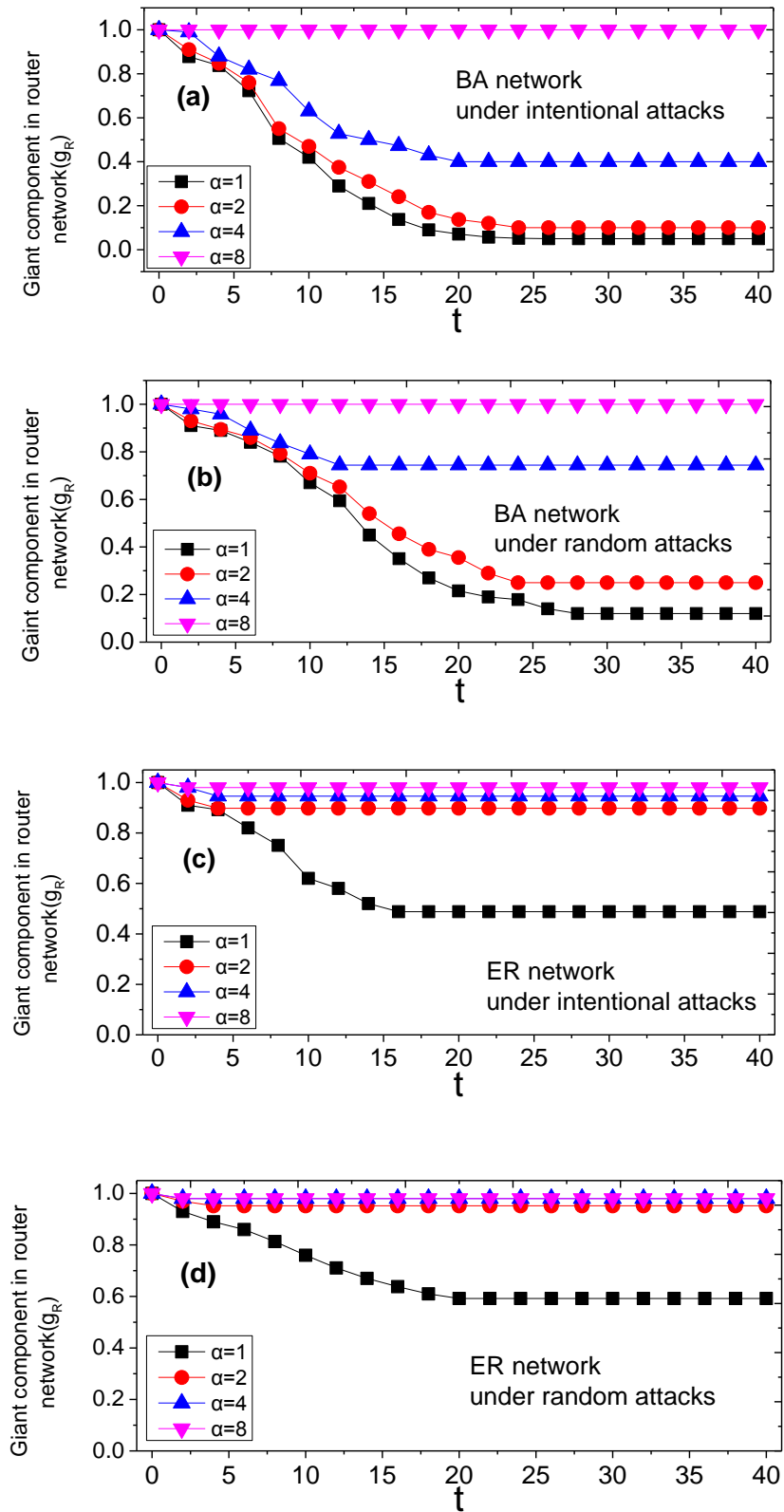


图 5-2 路由器网络最大连通片随时间的变化

其中, $l_0 = 10$ (Gbps) 是链路的初始传输容量。 $\beta = 0.1$ (s) 来衡量数据包的平均往返时间 (RTT)。

业务设置: 100,000 条业务随机分布在网络中, 即源节点、目的节点随机选择, 并依据 Dijkstra 最短路径算法分配路径。业务随机赋予优先级 $\varphi = 1, 2, 3$, 业务带宽需求服从正态分布 $D \sim N(0.2, 0.4)$ 。

故障触发: 随机选择或蓄意选择 5% 的路由器进行攻击触发级联故障, 使业务重新选路, 带宽资源重新分配。受到攻击的路由器和相连接的链路将在网络拓扑中删除。随机攻击时, 路由器随机移除; 蓄意攻击时, 具有更高度数的路由器被移除。

5.3.2 网络级联失效

图 5-2 显示了路由器网络在级联失效过程中最大连通片 g_R 的变化过程。横坐标 t 是级联失效中业务负载重新分配的次数, 表示了网络失效的时间进程。当业务再分配时, 重新计算 g_R 值。

对于 BA 网络, 如图 5-2 (a) 和 (b) 所示, 在蓄意攻击和随机攻击下都十分脆弱。蓄意攻击下 g_R 值与随机攻击相比平均下降 41%。在低链路容量环境下 ($\alpha = 1, 2$), g_R 下降迅速, 基本在 $t \leq 20$ 时, g_R 已下降至 0, 网络基本瓦解, 仅剩孤立节点。在高链路容量环境下 ($\alpha = 4, 8$), g_R 下降速度有所减缓, 网络最大连通片可以在级联失效后期 ($t \geq 20$) 维持在 $0.4 \leq g_R \leq 1$, 有一半以上的网络节点和链路处于最大连通片中具备传输能力。对于 ER 网络, 如图 5-2 (c) 和 (d) 所示, 在蓄意攻击和随机攻击下都有较高的鲁棒性。仅在低链路容量环境下 ($\alpha = 1$), g_R 有下降趋势, 但仍能保持 $g_R \geq 0.4$ 。在高链路容量环境下 ($\alpha = 4, 8$), 可以维持更高的值 $0.8 \leq g_R \leq 1$ 。这意味着绝大部分的网络节点和链路仍能正常工作。

此外, 可以看出增加链路容量可以有效提高网络抵御流过载引发的级联失效。以 BA 网络受到蓄意攻击为例 (图 5-2 (a)), 随着链路容量的增加 ($\alpha = 1, 2, 4$), g_R 由 0 增加至 0.1、0.4。当链路传输容量足够大时 ($\alpha = 8$), 故障几乎不对网络造成传播影响。

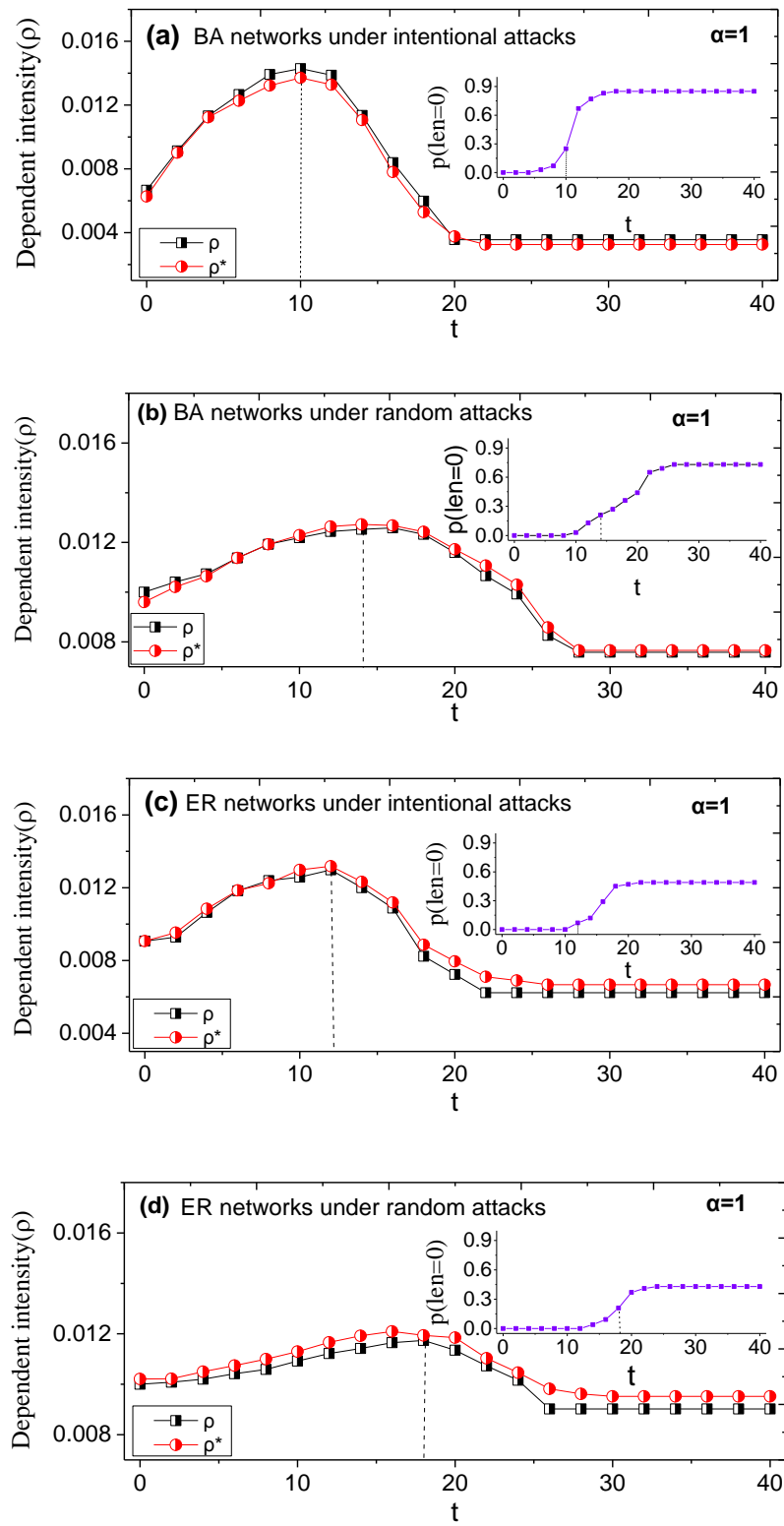


图 5-3 业务流对网络的依赖强度。子图显示了故障过程中传输中断的业务流比例。

5.3.3 业务依赖强度（传输效率）

图 5-3 显示了 $\alpha=1$ 时不同攻击类型下，业务流对网络依赖强度 ρ 的变化。

从图 5-3 可以看到，依赖强度 ρ 和理论值 ρ^* 具有大致相同的趋势，表明 ρ 与业务的平均路径长度确实具有正相关关系 $\rho \propto \bar{s}$ ，因此， ρ 可以反映业务的传输效率（负相关）。当 $t=0$ 时，即网络尚未受到攻击时，相比于 ER 网络，BA 网络具有较低的依赖强度 ρ 值，因此 BA 网络比 ER 网络在数据传送方面效率更高。

故障初期，初始失效的节点和链路从网络中移除，业务重新分配并选择其他次优路径。由于故障过程中网络结构越来越稀疏，业务流新路径通常较之前会更长。因此，依赖强度 ρ 随着故障的蔓延出现初步增加的趋势，相对地，业务的传输效率也变得越来越低。此外，依赖强度 ρ 上升期的长短也反映了网络恢复力和承受故障能力的高低。在 ρ 上升期间，尽管业务的传输效率下降，但网络依然具有较高的连通性，仍能成功建立业务传输的备份路径。 ρ 上升期越长，网络的抗毁性越好。BA 网络中 ρ 上升期（ $0 \leq t \leq 10$ 和 $0 \leq t \leq 14$ ）比 ER 网络（ $0 \leq t \leq 12$ 和 $0 \leq t \leq 18$ ）短，这意味着网络发生级联失效时，BA 网络的抗毁性较差。

随着故障的进一步蔓延， ρ 开始减少。为解释这一现象，统计了故障过程中传输失败的业务比例（图 5-3 中子图），传输失败的业务流其传输路径长度为 0。 ρ 从增加到减少的转折点恰是网络中开始有一定数量的业务未能成功选路，从而导致 ρ 和 ρ^* 的减少。此外， ρ 下降趋势表明网络的连通性由良好逐步转向恶化。蓄意攻击下的 BA 耦合网络中（图 5-3（a））， ρ 值的下降点为 $t=10$ ，此时网络最大连通片的比例为 $g_R \approx 0.4$ （图 5-2（a））；当 $t=20$ 时， ρ 值下降近 50%，近 90% 的业务流传输失败，此时网络最大连通片的比例仅约为 $g_R \approx 0.1$ （图 5-2(a)）。

此外，不同网络结构中 ρ 的纵向幅值变化也不同。BA 网络 ρ 幅值变化范围较大，表明故障对于 BA 网络中的业务重路由路径长度影响更大，尤其是蓄意攻击时。对于 ER 网络， ρ 幅值变化范围较小，表明故障对于 ER 网络中业务路径长度影响较小。因此，ER 网络相较于 BA 网络业务传输稳定性更高。

5.3.4 业务竞争范围

图 5-4 显示了故障过程中业务流耦合关联网络最大连通片 g_F 的波动情况，同时 g_F 也反映了业务最大竞争范围的变化情况。每当业务重新分配时，业务之间的竞争关系重建建立，重新计算 g_F 值。

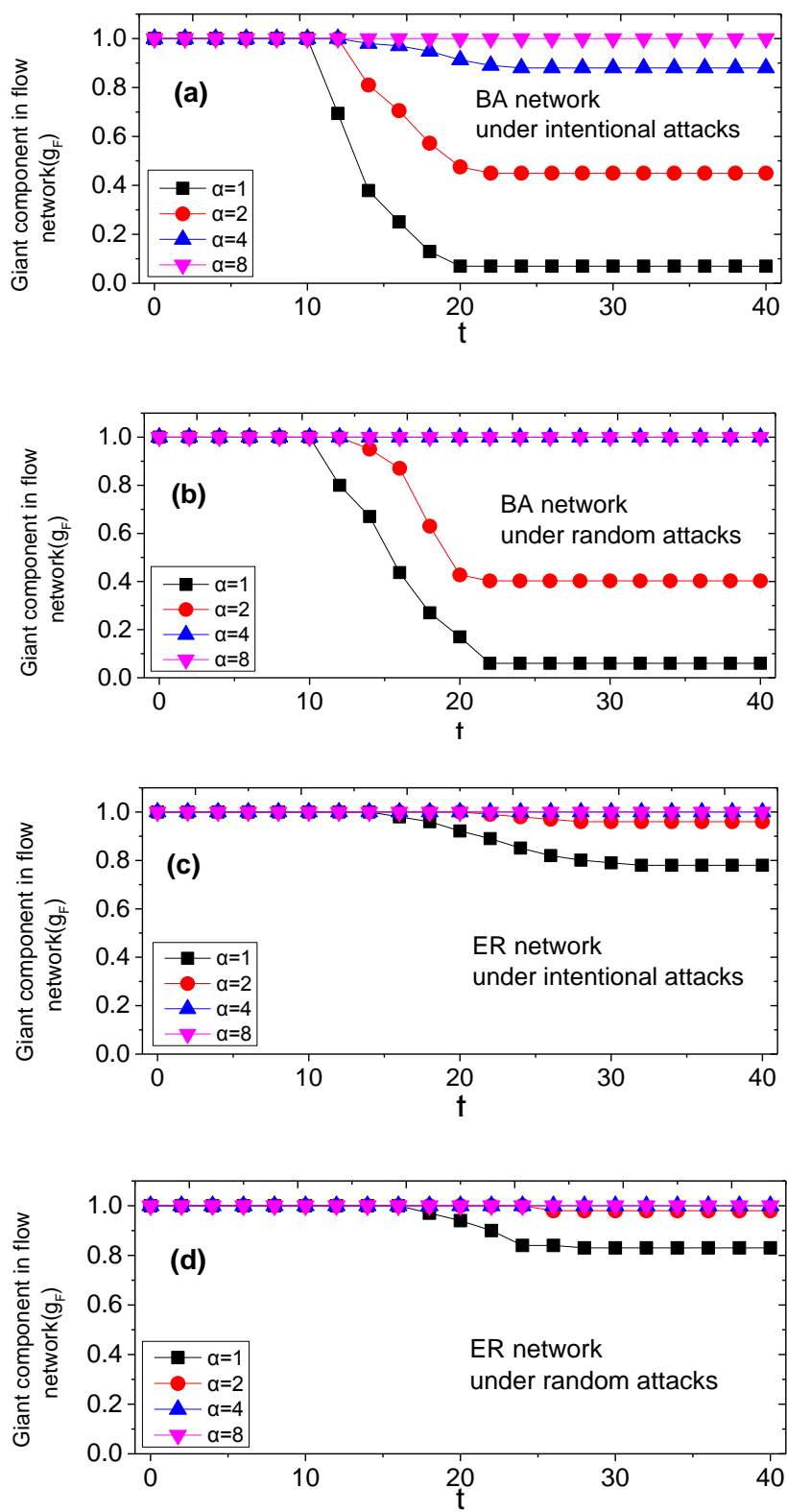


图 5-4 业务最大竞争范围

从图 5-4 (a) 和 (b) 观察得出, BA 网络中业务的竞争关系更易受到网络故障的影响。在未受攻击时 ($t=0$), BA 网络中业务最大的竞争区域为 $g_F=1$, 由于业务的重路由功能, 竞争区域在网络级联失效初期保持不变。但随着级联故障的蔓延, g_F 开始迅速下降。这是由于故障后期, 路由器网络的连通性逐渐恶化(图 5-2), 使得许多业务难以成功找到备用路径, 从而变得与其他业务孤立(图 5-3 中子图), 失去了竞争网络资源的能力, 此时业务最大竞争区域迅速下降。

在 ER 网络中(图 5-4 (c)、(d)), 无论是随机攻击还是蓄意攻击, g_F 数值变化并不明显, 大多数业务都能够维持竞争网络资源的能力。仿真结果显示, ER 网络的业务最大竞争区域是 BA 网络的 4 倍以上。此外, 通过增加链路容量(调节 α 值), 业务最大连通片的比例也将增大, 可以明显地提高网络鲁棒性, 更多的业务能够在级联失效中找到备份路径成功传输。

5.3.5 业务竞争强度

图 5-5 显示了不同攻击类型下, 业务竞争强度 γ 的变化。可以观察到, 在级联故障初期, 竞争强度 γ 不断增加。这段时期大部分受失效链路影响的业务能够重新路由, 而随着链路数量和网络资源的减少, 业务连接度 k 不断增加, 传输带宽 b 不断下降, 共同导致了竞争强度 γ 的增加(公式 5-25)。随着故障继续扩散, γ 开始下降。这是由于随着网络连通度的恶化, 越来越多的业务传输失败。对于传输失败的业务, 其连接度 $k=0$ 、传输带宽 $b=0$, 导致了业务平均竞争强度的下降。

业务的竞争强度受业务连接度 k 和业务传输带宽 b 两个因素的共同影响, 忽略后者的影响, 可以得到竞争强度的最大值 γ_{\max} (公式 5-26)。从图 5-5 可以看出, γ_{\max} 趋势和 γ 基本一致, 且随着故障的持续, 两者差异逐渐变小(图 5-5 中子图)。由于 γ_{\max} 的计算复杂度较小, 因此可以用来粗略评估业务竞争强度的大小。

最后, 由于 BA 网络具有明显的异构特性, 业务的传输路径更为集中, 业务连接度 k 更易受网络攻击的影响。在此情况下, 与 ER 网络相比, BA 网络中业务的竞争强度受网络攻击的影响更为明显。

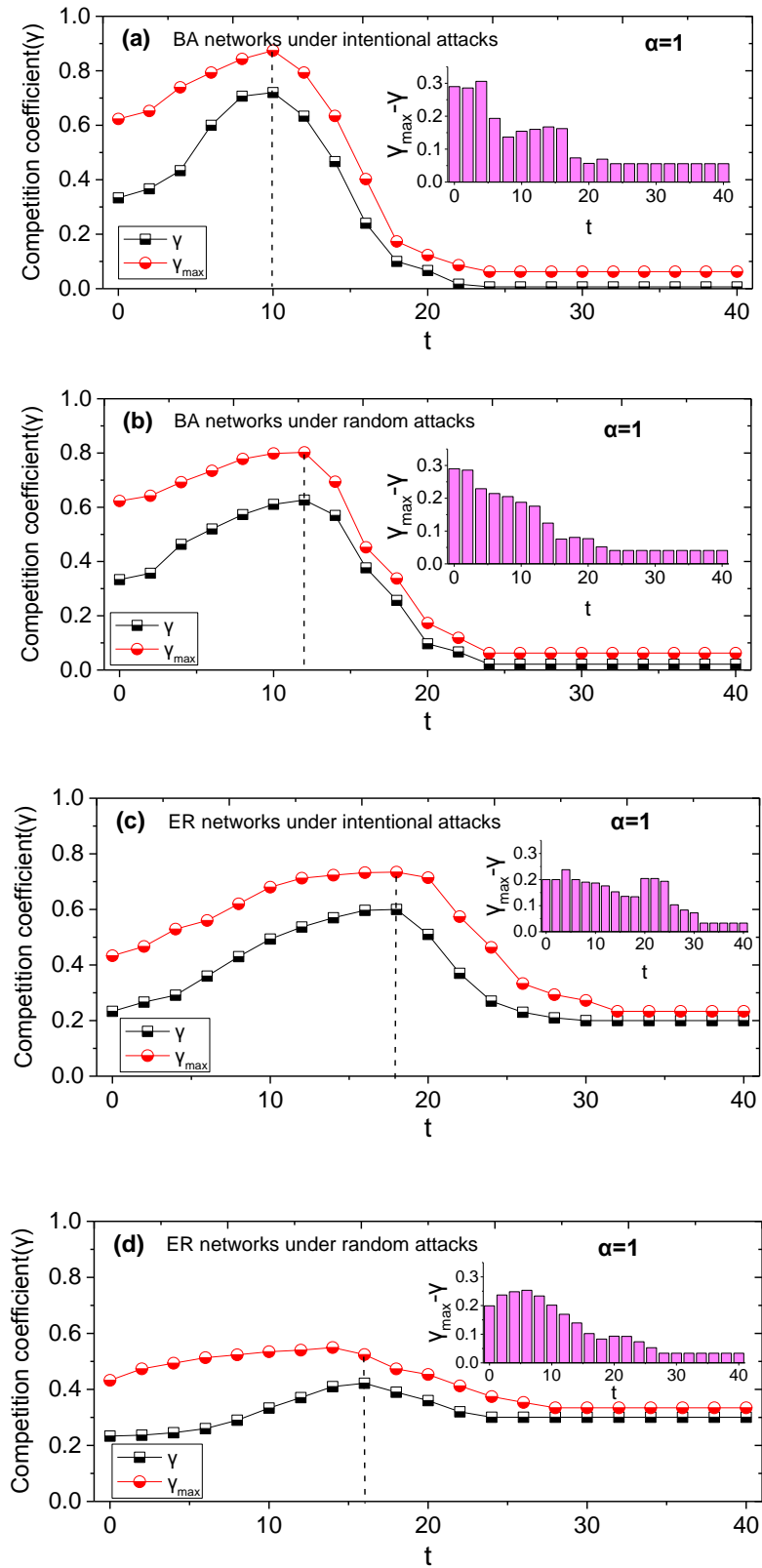


图 5-5 业务竞争强度。子图为仿真值与理论最大值的偏差。

5.4 本章小结

本章建立了业务层/网络层的双层网络模型，基于该模型主要考察链路级联失效过程中业务的动态行为变化，包括业务传输行为和业务间的竞争行为。主要得到以下结论：1) 业务流对网络的依赖强度与业务的平均最短路径长度成正比，可间接反映业务传输效率和网络传输性能的高低。故障初期，由于业务的重路由路径通常较之前更长，因此依赖强度逐渐增加；随着故障蔓延，越来越多的业务传输失败，此时传输依赖强度迅速下降；2) 业务最大竞争区域在故障初期保持稳定，但是随着网络连通度的恶化，业务间竞争关系产生隔离，竞争区域急剧下降；3) 故障初期，由于业务连接度和拥塞程度逐步增加，业务竞争强度越来越大；随着故障传播，大面积业务传输失败，网络业务平均竞争程度逐步降低。由于无标度网络中业务分布的不均匀性，业务竞争强度受级联故障的影响更明显。

参考文献

- [1] Korkali, M., Veneman, J. G., Tivnan, B. F., et al. Reducing Cascading Failure Risk by Increasing Infrastructure Network Interdependence[J]. Scientific Reports, 2017, 7:44499.
- [2] Chen, S. M., Xu, Y. F., Nie, S. Robustness of network controllability in cascading failure[J]. Physica A Statistical Mechanics & Its Applications, 2016, 471:536-539.
- [3] Zhao, D., Wang, Z., Xiao, G., et al. The robustness of interdependent networks under the interplay between cascading failures and virus propagation[J]. EPL, 2016, 115(5): 58004.
- [4] Hong, S., Lv, C., Zhao, T., et al. Cascading failure analysis and restoration strategy in an interdependent network[J]. Journal of Physics A Mathematical & Theoretical, 2016, 49(19):195101.
- [5] Havlin, S. Robustness of a Network of Networks[J]. Physical Review Letters, 2010, 107(19):195701.
- [6] Cai, Y., Li, Y., Cao, Y., et al. Modeling and impact analysis of interdependent characteristics on cascading failures in smart grids[J]. International Journal of Electrical Power & Energy Systems, 2017, 89:106-114.

- [7] Havlin, S., Kenett, D. Y. Cascading Failures in Interdependent Economic Networks[A].// Proceedings of the International Conference on Social Modeling and Simulation, plus Econophysics Colloquium [C]. Heidelberg: Springer International Publishing, 2015:87-97.
- [8] Zhou, D., Bashan, A., Berezin Y, et al. On the Dynamics of Cascading Failures in Interdependent Networks[J]. 2012, 1(1):166-169.
- [9] Rohden, M., Jung, D., Tamrakar, S., et al. Cascading failures in ac electricity grids[J]. Physical Review E, 2016, 94(3-1):032209.
- [10] La, R. J. Cascading Failures in Interdependent Systems: Impact of Degree Variability and Dependence[J]. IEEE Transactions on Network Science & Engineering, 2017, 99:1-1.
- [11] Dey, P., Mehra, R., Kazi, F., et al. Impact of Topology on the Propagation of Cascading Failure in Power Grid[J]. IEEE Transactions on Smart Grid, 2016, 7(4):1970-1978.
- [12] Haseltine, C., Eman, E. S. Prediction of Power Grid Failure Using Neural Network Learning[A].// IEEE International Conference on Machine Learning and Applications[C]. USA: IEEE, 2017:505-510.
- [13] Zhang, P., Cheng, B., Zhao, Z., et al. The resilience of interdependent transportation networks under targeted attack[J]. 2013, 103(6):613-616.
- [14] Tan, F., Xia, Y., Zhang, W., et al. Cascading failures of loads in interconnected networks under intentional attack[J]. EPL, 2013, 102(2):28009.
- [15] Cai, Y., Cao, Y., Li, Y., et al. Cascading Failure Analysis Considering Interaction Between Power Grids and Communication Networks[J]. IEEE Transactions on Smart Grid, 2016, 7(1):530-538.
- [16] Lee, E. J., Goh, K. I., Kahng, B., et al. Robustness of the avalanche dynamics in data-packet transport on scale-free networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 71(5):056108.
- [17] Hsu, C. J., Freeh, V. W., Villanustre, F. Trilogy: Data placement to improve performance and robustness of cloud computing[A].// IEEE International

- Conference on Big Data[C]. Boston: IEEE, 2018:2442-2451.
- [18] Sanguankotchakorn, T., Manguin, F. Performance improvement of a random access channel protocol over M2M communications[A].// International Congress on Ultra Modern Telecommunications and Control Systems and Workshops[C]. Munich: IEEE, 2017:67-72.
- [19] Crucitti, P., Latora, V., Marchiori, M. Model for cascading failures in complex networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 69(4 Pt 2):045104.
- [20] Guo, C., Wang, L., Zhou, F., et al. Analysis on the “Robust Yet Fragile” Nature of Internet: Load, Capacity and the Cascading Failure Avalanche Effect[A].// International Conference for Young Computer Scientists [C]. Hunan: IEEE, 2008:2149-2153.
- [21] Doyle, J. C., Alderson, D. L., Li, L., et al. The "robust yet fragile" nature of the Internet[J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(41):14497-502.
- [22] Tan, F., Xia, Y., Wei, Z. Robust-yet-fragile nature of interdependent networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2014, 91(5):052809.
- [23] Liu, X., Ji, Y. The effects of vendor-specific router implementation on convergence[A].// International Conference on Communications and NETWORKING in China [C]. Beijing: IEEE, 2006:1-5.
- [24] Callaway, D. S., Newman, M. E., Strogatz, S. H., et al. Network robustness and fragility: percolation on random graphs[J]. Physical Review Letters, 2000, 85(25):5468-5471.
- [25] Su, Z., Li, L., Peng, H., et al. Robustness of interrelated traffic networks to cascading failures[J]. Sci Rep, 2014, 4:05413.
- [26] Iranmanesh, S. A novel queue management policy for delay-tolerant networks[J]. Eurasip Journal on Wireless Communications & Networking, 2016, 2016(1):1-23.

- [27] Zhao, H., Niu, W., Qin, Y., et al. Traffic Load-Based Dynamic Bandwidth Allocation for Balancing the Packet Loss in DiffServ Network[J]. Journal of Network New Media, 2012:99-104.
- [28] He, R., Xie, H., Fang, X. A long-term proportional fair dynamic bandwidth allocation scheme for EPON[A].// Wireless and Optical Communication Conference[C]. Chengdu: IEEE, 2016:1-5.
- [29] Liu, K., Lee, J. Y. B. On Improving TCP Performance over Mobile Data Networks[J]. IEEE Transactions on Mobile Computing, 2016, 15(10):2522-2536.
- [30] Hong, S., Lv, C., Zhao, T., et al. Cascading failure analysis and restoration strategy in an interdependent network[J]. Journal of Physics A Mathematical & Theoretical, 2016, 49(19):195101.
- [31] Cai, Y., Cao, Y., Li, Y., et al. Cascading Failure Analysis Considering Interaction Between Power Grids and Communication Networks[J]. IEEE Transactions on Smart Grid, 2016, 7(1):530-538.
- [32] Liu, R. R., Ming, L., Jia, C. X. Cascading failures in coupled networks: The critical role of node-coupling strength across networks[J]. Scientific Reports, 2016, 6:35352.
- [33] Zhu, G., Wang, X., Tian, M., et al. Cascading failures of interdependent modular small-world networks[J]. Modern Physics Letters B, 2016, 30(18):1650174.
- [34] Yong, Z., Lei, J., Xiao, J. W. Research on cascading failure in multilayer network with different coupling preference[J]. International Journal of Modern Physics C, 2017, 28(4):1750050.

第六章 总结与展望

通过对业务的流量特征、流量分布、服务质量等方面进行数学建模,网络业务流行为可以应用于网络资源优化分配、业务服务质量提升、网络整体规划、网络安全等方面^[1-3]。现代通信网络的发展日渐复杂化^[4-5],与此同时,网络用户数目也与日俱增,尤其是互联网时代的到来,使得网络业务类型更加复杂多元,业务突发性和实时性也逐渐增强^[6-7]。如何在大规模的复杂网络环境中,捕捉和刻画网络业务行为,使之服务于网络整体运行,成为迫切需要解决的问题。

本文针对已有研究不足,基于复杂网络和弹性力学相关理论知识分析业务与业务、业务与网元间微观与宏观的耦合关联,研究不同网络环境下(业务突变与网络故障),全网业务的动态耦合特性、波动传播特性与相应控制策略。

6.1 主要研究成果

(1) 业务流耦合关联网络的小世界特性

本文依据业务之间共享资源的关联关系,构建了无向加权的业务流耦合关联网络,其中节点表示业务流,边表示业务之间的关联关系。通过对业务流耦合关联网络的结构特征进行研究,发现业务流在网络中随机分布并达到一定数目时具有“小世界效应”:平均路径长度小、聚集系数高。具体表现为业务之间相互影响的平均路径较短且业务的聚集程度较高,体现出网络业务之间密切的关联性,并且有助于理解业务间耦合作用的扩散特性。此外,通过对业务的度分布进行统计,发现了网络中“超级流”的存在,超级流通常具有非常高的度数,且影响范围极广。

(2) 评价网络拥塞的新指标

本文基于通信系统与弹性物理系统的相似性,建立了业务流弹性物理模型,并利用弹性势能概念定义了网络“弹性拥塞系数”。实验结果发现,“弹性拥塞系数”不仅可以反映网络拥塞情况,而且可以区别相同优先级业务和比例优先级业务之间的网络拥塞差异。此外,业务弹性拥塞系数的概率分布在网络拥塞环境下大致满足幂律分布,表明网络中大部分业务的拥塞程度较低,仅有少数业务处于

严重拥塞状态。

(3) 突变业务流的影响传播特征

本文建立了突变业务流对全网业务服务质量的影响传播模型,使用“邻居阶数”描述业务的邻居关系,并定义“耦合概率”描述相邻两阶邻居业务的耦合作用。通过仿真分析,突变业务流的传播行为有如下特征:1) 突变业务流的影响传播是一个随着邻居阶数增长而大幅下降的过程,且对奇数阶邻居和偶数阶邻居的影响正负交替进行;2) 随着网络拥塞程度的增加,突变业务流的传播强度也显著增加,尤其是突变流对其一阶邻居的影响强度;3) 当网络中业务数目固定时,网络规模的扩大可以降低突变业务流对全网业务的平均影响强度;4) 在相同网络规模和随机业务分布情况下,小世界网络中业务之间的耦合概率较低,在降低突变业务流影响强度方面优于随机网络和无标度网络。

(4) 突变业务流影响传播免疫控制策略

本文从病毒免疫角度和业务社团属性出发,提出了一种基于业务社团结构的免疫控制策略,有效控制了突变业务流造成的全网业务带宽波动现象。通过将业务划分为不同等级的社团并免疫社团边界的关键业务,有效阻止了业务社团间的影响传播,最终将突变流的影响控制在极小范围之内。仿真结果表明,对于业务流耦合关联网络,基于社团的免疫控制策略优于随机免疫策略和目标免疫策略。

(5) 网络故障过程中业务流的动态行为特征

本文建立了业务层/网络层双层网络,基于该网络,考察了链路级联失效过程中,业务的动态行为变化,包括业务传输行为和业务间的竞争行为。主要得到以下结论:1) 业务流与网络之间的依赖强度与业务平均路径长度成正比,可以间接反映业务的传输效率;2) 业务最大竞争区域反映了业务保持网络资源占有能力的高低,在故障初期保持稳定,但随着后期网络连通度的恶化,竞争区域急剧下降;3) 业务竞争强度与业务度分布和拥塞程度相关,随着故障的传播,出现先增后减的趋势。此外,由于无标度网络中业务分布的不均匀性,业务竞争强度受级联故障的影响更明显。

6.2 主要创新点

本文围绕网络业务流行为进行了理论建模和仿真分析,不仅从通信网络本身出发,并结合复杂网络和弹性物理学的相关理论知识,挖掘业务流行为特征。本

文主要有以下创新点:

(1) 提出了一种基于弹性力学的业务流耦合关联网络模型, 挖掘业务服务质量间的耦合机理。

本文将弹性物理系统映射到通信系统中, 构建了业务流服务质量间的弹性物理耦合关系, 刻画了网络运行过程中业务呈现出的微观耦合特性。进一步提出将业务流弹性耦合关系抽象成逻辑上关联网络的方法, 构建了业务流耦合关联网络, 一方面可以直观反映任意业务流与其他流的逻辑影响关系, 另一方面可以明确业务流之间影响传播的扩散特点。利用复杂网络相关分析方法, 对业务流耦合关联网络的业务聚集特性、度分布等方面进行规律挖掘, 发现了业务流“小世界效应”的存在。

(2) 提出了一种突变业务流影响传播模型, 分析其传播机理与传播规律。

本文提出了一种基于耦合概率与邻居关系的突变业务流影响传播模型, 模型精确度随着网络拥塞程度的增加而不断提高。利用该模型分析了不同业务突变强度、不同网络拥塞程度、不同网络结构类型下突变流的传播范围、传播强度等规律。仿真结果表明, 突变流的传播强度会随着业务邻居阶数增长而大幅下降, 下降幅度与各阶业务的数目相关, 且对奇数阶邻居与偶数阶邻居的带宽影响相反。

(3) 提出了一种基于业务社团结构的免疫控制算法, 保障业务安全可靠传输。

传统突变流影响控制策略多集中于增加系统开销改变业务流的发送速率与强度, 一方面降低了业务的传输效率, 另一方面传播影响路径并没有被切断。本文引入传染病模型中“免疫”的概念, 将业务流耦合关联网络中关键节点进行“免疫”预处理, 在不影响大部分业务正常传输的情况下, 切断突变流的传播途径。在关键业务节点的选择上, 通过对业务流聚集特点的研究, 发现了业务间的社团属性。社团边界节点通常对影响传播路径具有关键作用, 因此, 本文将业务进行两级社团划分, 对社团边界业务流进行“免疫”处理, 有效地控制了突变流在社团间的影响传播。

(4) 提出了一种基于双层网络的级联故障模型, 分析业务流动态行为特征。

双层网络模型通常用来表述两种类型网络间的关联关系, 应用于网络可靠性与鲁棒性研究, 如电力网络与通信网络、地面交通网络与航空网络。不同于传统

双层网络模型，本文建立的业务层/网络层双层网络模型中业务层为逻辑网络，并借此建立了基于流过载的网络链路级联失效模型。该模型可以清晰地反映在网络故障发生时业务流与网络之间依赖关系的动态变化，并有助于分析网络级联失效时业务流传输行为、竞争行为的内在原因和机理。

6.3 未来研究方向

本文对大规模复杂网络中业务流的耦合行为、动力学传播行为进行了研究。结合本文研究内容，未来的研究工作拟从以下两方面入手：

(1) 探究业务层/网络层双层网络的渗流相变点

渗流相变^[8]体现了网络由一个完整的功能系统逐渐变为离散孤立节点的过程，随着攻击节点比例的增加，网络一般表现出连续相变或非连续相变两种特性。对渗流阈值（相变点）的研究可以帮助人们预知网络由鲁棒向脆弱转变的临界条件，进而做出及时的补救和恢复措施，具有实际应用价值。下一步拟研究节点攻击比例连续变化情况下，业务层和网络层最大连通片的变化情况，探究渗流相变点。

(2) 网络级联失效模型中增加网络恢复机制

随着通信网络智能化的逐步发展，网络故障的预判预测已成为网络安全发展的必然趋势^[9]。在遭遇攻击时，网络会在一定时间内进行有序恢复^[10]，网络故障的传播通常伴随着网络的恢复。本文目前建立的级联故障模型尚未考虑网络节点和链路的恢复机制，下一步拟将故障传播和网络恢复两个过程进行叠加，探究此时网络业务体现出的复杂行为特性。

参考文献

- [1] Xu, K., Zhang, Z. L., Bhattacharyya, S. Profiling internet backbone traffic: behavior models and applications[A].// Conference on Applications[C]. Philadelphia: ACM, 2005:169-180.
- [2] Li, K., Xu, X., Swamy, M. Modelling and Analysis of Regional Service Behavior Properties of Mobile Internet Applications[J]. IEEE Access, 2017, 5: 4795 - 4807.
- [3] Grandemange, Q., Ferveur, O., Gilson, M., et al. A live network AS-level traffic

- characterization[A].// International Conference on Computing, Networking and Communications[C]. Santa Clara: IEEE, 2017:901-905.
- [4] Wu, D., Zhang, X. On Study of Construction of New Generation Intelligent Communication Network for Distribution System[J]. 2017, 242(1):012091.
- [5] Pérez-Romero, J., Sallent, O., Ferrús, R., et al. Artificial Intelligence-based 5G network capacity planning and operation[A].// International Symposium on Wireless Communication Systems[C]. Brussels: IEEE, 2016:246-250.
- [6] Kim, J., Kim, S., Bang, J., et al. Adaptive Mode Selection in D2D Communications Considering the Bursty Traffic Model[J]. IEEE Communications Letters, 2016, 20(4):712-715.
- [7] Yang, Y., Luo, H., Xu, H., et al. Towards Real-Time Traffic Sign Detection and Classification[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(7):2022-2031.
- [8] Gao, J., Buldyrev, S. V., Stanley, H. E., et al. Percolation of a general network of networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2013, 88(6):062816.
- [9] Duenas, J. C., Navarro, J. M., Hugo, A. P. G., et al. Applying Event Stream Processing to Network Online Failure Prediction[J]. IEEE Communications Magazine, 2018, 56(1):166-170.
- [10] Xie, K., Peng, C., Wang, X., et al. Accurate recovery of internet traffic data under dynamic measurements[A].// IEEE Conference on Computer Communications[C]. Atlanta: IEEE, 2017:1-9.

附录 1 缩略语

BA	Barabási-Albert Network	BA 无标度网络
BGLL	Blondel-Guillaume-Lambiotte-Lefebvre	层次聚类社团算法
CIR	Committed Information Rate	保证信息速率
ER	Erdős-Rényi Network	ER 随机网络
FIN	Flow Interaction Network	业务流耦合关联网络
ICMP	Internet Control Message Protocol	控制报文协议
IP	Internet Protocol	互联网协议
LSP	Label Switching Path	标签交换路径
NFV	Network Function Virtualization	网络功能虚拟化
NW	Newman-Watts Network	NW 小世界网络
PIR	Peak Information Rate	峰值信息速率
PW	Pseudo Wire	伪线
RD	Random Network	随机网络
RYF	Robustness-Yet-Fragile	鲁棒脆弱性
SDN	Software-Defined Network	软件定义网络
SF	Scale-Free Network	无标度网络
TCP	Transmission Control Protocol	传输控制协议
UDP	User Datagram Protocol	用户数据报协议
WS	Watts-Strogatz Network	WS 小世界网络
WWW	World-Wide-Web	万维网
4G	Fourth Generation	第四代移动通信

致谢

随着博士论文的完成，心中不禁百感交集。最初，怀着无限的憧憬开始博士生涯，如今这份美好的愿望即将成真。在此，我要深深地感谢攻读博士学位期间所有在学习和生活上给予过关心和帮助的老师、同学、朋友和亲人，是你们的鼓励和陪伴，让我的科研道路走得坚定不移。

首先，要感谢我的导师纪越峰教授。您引领我步入了学术的殿堂，并在科研上对我悉心指导。您敏锐的科研洞察力、严谨的治学作风和孜孜以求的敬业精神深深地影响着我，鞭策我不断前进，为我树立了科研的榜样。此外，您不仅为我创造了优越的科研环境，还提供机会鼓励我到国外进行访问学习，使我在学术道路上快速成长。再次向您表示由衷的感谢。

其次，要感谢在学术和生活上给予耐心指导的顾仁涛副教授。您花费大量的时间与精力与我讨论学术问题，几乎每篇论文和每个项目都凝聚了您的心血，使我在科研方法、研究思路和论文写作方面受益颇深。您丰富的学术经验、严谨的工作态度和永远积极奋进的科研精神使我敬佩不已。感谢我的博士生涯有您的帮助与陪伴。

同时，感谢美国联合培养期间在学术上给予指导的 H. Eugene Stanley 教授。您在复杂网络领域和经济物理学领域卓越的研究成果让我的敬仰之心油然而生。与您合作的论文得到了您细心的指导与修正，使我的论文写作水平进一步提升。感谢您在一年时间里对我的帮助和鼓励。

感谢一起为科研努力奋斗的小伙伴们。感谢王鑫师姐、张士宗师兄、金伟祺师兄在科研上对我的帮助；感谢谭艳霞师妹、王刚师弟、邢真师妹、郑豪豪师弟、张一帆师弟、刘智辉师弟在科研道路上的陪伴；感谢我的博士室友，王静贤、高雅婷、葛梦舒；感谢国外生活一直陪伴左右的鲁韵帆、赵钦小伙伴。由衷的感谢你们。

衷心地感谢我的父母，是你们从小给予我良好的教育熏陶，让我能在科研的道路上勇敢前行！无论我在科研期间遇到任何困难，你们都能理解和鼓励我，让我能积极乐观地勇往无前！

最后，感谢专家评委们在百忙之中抽出时间审阅论文，并提出宝贵的意见。

武晓宇

2018 年 4 月于北京邮电大学

攻读学位期间发表的学术论文目录

期刊论文

- [1] Wu X., Gu R., et.al . Flow interaction based propagation model and bursty influence behavior analysis of Internet flows[J]. Physica A: Statistical Mechanics & Its Applications, 2016, 462:341-349. (SCI Indexed, WOS: 000381841900031, IF: 2.243)
- [2] Wu X., Gu R. Ji Y. Bandwidth turbulence control based on flow community structure in the Internet[J]. Europhysics Letters, 2016, 116(1):18005. (SCI Indexed, WOS: 000388369100029, IF: 1.975)
- [3] Wu X., Gu R. Ji Y. Structural Modeling and Characteristics Analysis of Flow Interaction Networks in the Internet[J]. Chinese Physics Letters, 2015, 32(6): 068901. (SCI Indexed, WOS: 000358597200050, IF: 0.800)
- [4] 武晓宇,刘国军,等. 分组传送网承载电力多业务的隔离性与关联性分析[J]. 电力系统自动化, 2016, 40(20):140-146. (EI Indexed, 检索号: 20164302945936)

申请专利

- [1] 刘国军, 武晓宇等. 一种分组传送网业务隔离度的测试系统及测试方法[P]. 中国专利: 201511020420.7, 2017-07-07.
- [2] 顾仁涛, 武晓宇等. 基于分组业务流社团结构的服务质量波动传播控制方法[P]. 中国专利: 201611157782.5, 2017-03-22.