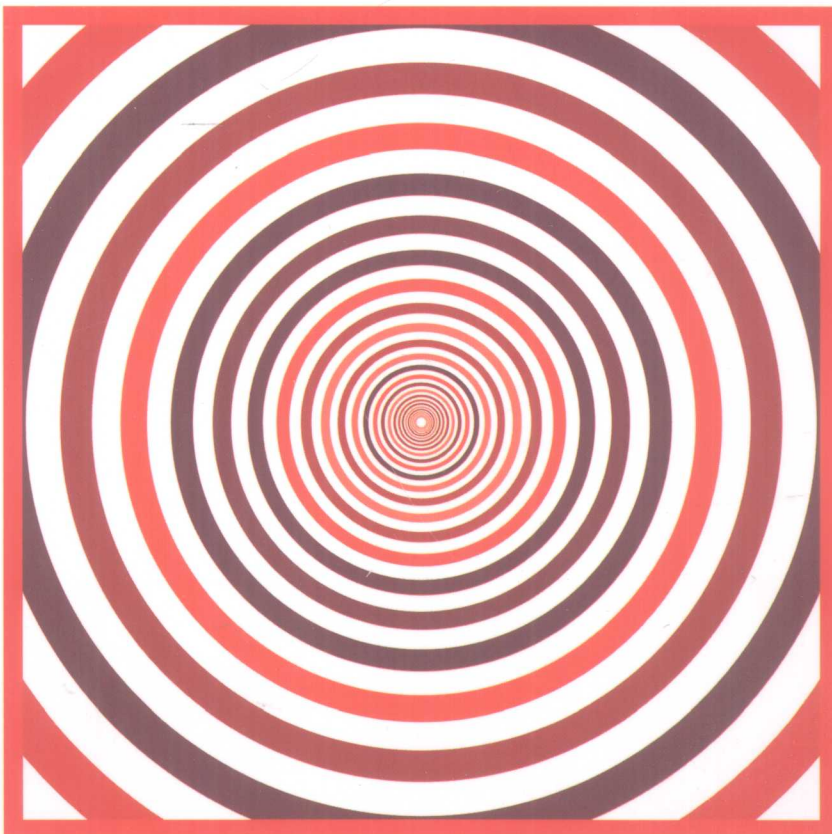


Network Information Theory

网络信息论

樊平毅 编著

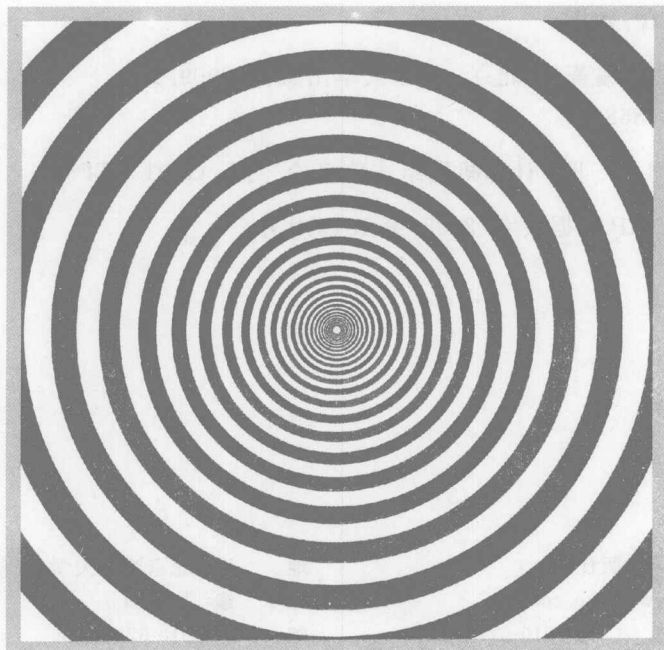
清华大学出版社



樊平毅 编著

网络信息论

Network Information Theory



清华大学出版社

北京

内 容 简 介

本书重点介绍了网络信息论研究的基本问题和从事网络信息论研究需要的基础知识,并结合目前网络信息论研究的热点课题,介绍了网络编码、数字喷泉编码以及跨层优化设计等。

本书主要内容包括网络信息论概论、传统的信道容量分析、图论与网络信息流、简单网络的信道容量、多址信道的容量分析、网络编码、数字喷泉技术以及通信网络的跨层优化设计理论等。通过对这些内容的介绍,使读者能更加深入地了解有关网络信息论的进展,并能开展相关方面的研究工作,推动我国在网络信息论方面的研究。

本书是现代网络信息论的入门教材,可作为高等院校电子信息学科高年级本科生和研究生相关课程的教材,也可供通信工程技术人员和科研人员作为参考书。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

网络信息论/樊平毅编著. —北京:清华大学出版社,2009.2

ISBN 978-7-302-18688-5

I. 网… II. 樊… III. 计算机网络—信息论 IV. G201 TP393

中国版本图书馆 CIP 数据核字(2008)第 154599 号

责任编辑:陈国新

责任校对:梁毅

责任印制:杨艳

出版发行:清华大学出版社

地 址:北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185×260 印 张:13 字 数:322 千字

版 次:2009 年 2 月第 1 版 印 次:2009 年 2 月第 1 次印刷

印 数:1~2500

定 价:25.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:030266-01

前言

自 2000 年后,网络信息论作为一个新的信息学科的发展方向,备受国际学术界的关注。在北美一些著名的大学,如 MIT、Stanford、UCSD 和加拿大的 Waterloo 等先后开设了网络信息论的研修课程,致力于介绍从事该研究所必需的理论基础和一些最新的研究成果,以推动整个信息学科向网络化、系统化和最优化发展,并使之更靠近工程应用的需求。

在这种科学研究逐步国际化的大背景下,我们自 2003 年开始关注这一国际学术动态,并及时跟进,开展了部分研究。在 2004 年、2005 年和 2006 年,作者所领导的研究组连续三年得到国家自然科学基金和国家 863 项目的支持,使得我们有机会开展高水平的前沿性课题研究,在网络编码、跨层设计以及网络容量估计等方面取得一系列国际领先的科研成果。在 2006 年和 2007 年,作者在清华大学电子工程系通信与信息系统专业硕士生和博士生开设的通信学科前沿课中,多次讲解“网络信息论的历史与发展”、“无线网络中网络编码理论”等。为进一步推动和培养从事该研究方向的专业人才,2007 年,作者在清华大学电子工程系为高年级的本科生开设了“网络信息论”的研究性课程。本教材是近几年这些讲稿经过细致化后所形成的版本,它包括我们对国际研究动态的部分调研资料和作者所领导的研究组的一些研究成果。我们希望本教材的出版可以起到抛砖引玉的作用,帮助我国从事该研究方向的学者、专家以及广大的研究生尽快熟悉该研究领域所需要的基础知识,了解国际研究动态,把握研究主流课题,找到合适的研究切入点,为提升我国在这一新的研究领域的国际地位奠定基础。

本书的特色主要体现在以下几个方面:

- (1) 从学科发展的角度,指出了网络信息论研究的三个重要问题。
- (2) 从应用的角度,通过大量的例子,并利用网络信息论的观点解释了现代通信理论发展中几个重要研究方向的形成,以及它们未来的发展方向。
- (3) 系统介绍了从事现代网络信息论研究的基础。
- (4) 在介绍传统网络信息论成果的基础上,介绍了现代网络信息论的几个重要发展方向和重要研究课题,如数字喷泉编码、网络编码以及跨层设计理论等。
- (5) 作为一本高级科普读物,本书从理论上解释了网络信息论在未来通信系统中的潜在价值。

本书的主要内容组织如下:

第 1 章介绍了网络信息论研究的三个基本问题,并对网络信息论的发展历史进行了简单的回顾,同时利用网络信息论的观点,对一些信息学科发展历史中里程碑性的研究课题进行了分析和解剖,意在展现网络信息论的研究价值。第 2 章和第 3 章是介绍网络信息论研究的基础知识。例如,第 2 章介绍了古典 Shannon 信息论以及信道容量计算理论。第 3 章介绍了图论的一些基础知识以及关于网络流的最大流最小割定理。第 4 章主要介绍了传统网络信息论有关简单网络的容量分析结果。第 5 章对多址接入网络的容量进行了归纳总

结。在第6章,从应用实现的角度介绍了网络编码的部分结果,和 Raymond Yeung 等编著的“网络编码”一书在内容上有很大的差别。在此,我们用了较多篇幅解释网络编码发展的多个研究层面。第7章介绍了数字喷泉编码技术,旨在推动我国这方面的研究进展。第8章介绍了跨层设计的几个重要研究方向,并通过一些例子介绍跨层设计思想的应用,最后对跨层优化设计未来的主要研究方向进行了归纳和总结。由于本学科的飞速发展,我们无法概括网络信息论涉及的所有内容,在此作者根据自己的观察将网络信息论涉及的基本问题予以简明的介绍,希望读者能通过阅读本书,对网络信息论的基本概貌有更深层次的了解。

借本书出版之际,向多年来一直关心我、培养我的博士生导师冯重熙教授致以衷心的感谢,他的指点和培养使我有机会接触到现代通信学科的研究精髓。感谢清华大学电子工程系,它为我提供了一个好的研究平台,使得我和我的研究组能够不断地学习与探索,保持与国际同行的交流与合作。感谢在“无线信息系统理论研究室”进行学习和从事科研的研究生们。感谢国家自然科学基金项目 No. 60472030、No. 60572085 和国家 863 项目 No. 2006AA01Z211 的支持。最后,感谢我的妻子和孩子的理解、支持和鼓励。

樊平毅

2008年8月于清华园

目 录

第 1 章 概 论	1
1.1 网络信息论研究的三个基本问题	1
1.2 网络信息论的发展历史	3
1.3 网络信息论的应用	4
1.4 网络信息论的未来发展方向	9
习题	11
参考文献	12
第 2 章 信道容量分析	13
2.1 信息熵理论	13
2.2 链式法则与信息处理	18
2.3 互信息的链式法则	19
2.4 离散信道的容量	20
2.4.1 离散信道的容量	20
2.4.2 信道容量的计算理论	22
2.5 微分熵与连续信道的容量公式	25
2.5.1 微分熵的定义	25
2.5.2 微分熵与离散熵的关系	26
2.5.3 相对熵与互信息	26
2.5.4 微分熵的极值问题	29
2.5.5 加性高斯噪声信道的信道容量公式	31
2.5.6 Shannon 信道容量公式	33
2.6 Fano 不等式与信道编码定理	35
2.7 信道容量的迭代算法	38
习题	41
参考文献	43
第 3 章 图论与网络信息流	44
3.1 图的基本概念	44
3.2 树图	45
3.2.1 树的基本性质	45
3.2.2 图的支撑(生成)树	46
3.2.3 最小支撑(生成)树及其算法	47

3.3 最短路径问题及算法	48
3.3.1 狄克斯拉(Dijkstra)算法	48
3.3.2 福特-摩尔-贝尔曼(Ford-Moore-Bellman)算法	50
3.4 网络最大流	51
3.4.1 基本概念与定义	51
3.4.2 最大流的计算——Ford-Fulkerson 算法	54
3.5 最小代价最大流问题	55
3.6 信息流理论介绍	56
习题	58
参考文献	60
第4章 简单网络的信道容量	61
4.1 引言	61
4.2 并行高斯信道的容量	61
4.3 有色高斯噪声信道的容量	63
4.4 基于反馈机制的高斯信道的容量	65
4.5 广播信道的容量	68
4.5.1 几个典型广播信道的例子	68
4.5.2 广播信道的一些基本定义	70
4.5.3 几个广播信道的容量域分析	72
4.5.4 改进的广播信道的容量界	75
4.6 接力通信系统的信道容量	75
4.7 非完全信息的接力模式的容量	77
4.8 合作模式的传输容量	79
4.9 有待研究的问题	82
习题	83
参考文献	84
第5章 多址信道的容量分析	85
5.1 引言	85
5.2 高斯多址信道的容量	85
5.3 中心控制模式的多址系统	88
5.3.1 固定用户发送功率的时分多址系统	88
5.3.2 固定系统内平均发送功率的时分多址系统	88
5.3.3 频分多址系统	89
5.3.4 码分多址系统	89
5.4 高斯多址信道容量的实现模式讨论	90
5.5 多信源编码问题	91
5.6 随机多址信道的容量分析	92

5.6.1	多址接入系统的模型	93
5.6.2	单时隙可实现的传输容量的定义	94
5.6.3	基本的编码与解码策略	94
5.6.4	公平准则下系统的最大可实现速率	97
5.7	随机多址衰落信道的容量分析	100
5.7.1	Rayleigh 信道模型的基本特征	100
5.7.2	可实现传输速率分析	101
5.7.3	公平准则下系统的最大可实现速率	104
5.8	未来的研究方向	106
	习题	107
	参考文献	108
第 6 章 网络编码		109
6.1	引言	109
6.2	网络编码的存在性	111
6.3	网络编码的增益	113
6.4	网络编码的构造方法	114
6.4.1	线性网络编码方法	114
6.4.2	代数型网络编码方法	117
6.4.3	随机网络编码方法	120
6.5	网络编码的复杂度分析	121
6.5.1	有线组播网络中编码节点的数目分析	121
6.5.2	无线自组织网络中编码节点数目分析	126
6.6	网络编码的实现问题	126
6.6.1	网络编码的容量性能分析	126
6.6.2	网络编码的延时模型	129
6.7	网络编码的机会调度策略	133
6.7.1	机会调度策略的原理	133
6.7.2	机会调度网络编码的性能分析	134
6.8	物理层网络编码	138
6.9	应用层网络编码	139
6.10	网络编码的应用	140
6.11	网络编码与业务保密	144
6.12	网络编码的发展趋势	146
	习题	147
	参考文献	148
第 7 章 数字喷泉技术		151
7.1	引言	151

7.2	数字喷泉码的设计策略	151
7.3	随机线性数字喷泉编码	153
7.4	LT 码	154
7.5	Raptor 码	158
7.6	信道的喷泉容量	159
7.7	Turbo 喷泉编码	161
7.7.1	并行 Turbo 喷泉	162
7.7.2	Multiple Turbo 喷泉	163
7.7.3	Turbo 喷泉的性能比较	164
7.8	数字喷泉技术的应用	165
7.9	数字喷泉码的发展	166
	习题	166
	参考文献	167
第 8 章	通信网络的跨层设计理论	168
8.1	网络的分层结构	168
8.1.1	应用层	168
8.1.2	表示层	169
8.1.3	会话层	169
8.1.4	传输层	169
8.1.5	网络层	170
8.1.6	数据链路层	171
8.1.7	物理层	172
8.2	网络的跨层设计模式介绍	172
8.3	跨层设计技术的应用	174
8.4	无线网络的信息容量	178
8.4.1	网络的传输容量	179
8.4.2	网络的信息流容量	179
8.4.3	网络的吞吐量	180
8.5	无线网络的路由选择与优化	180
8.5.1	纯路由处理技术	180
8.5.2	干扰控制路由处理技术	182
8.6	合作通信	185
8.7	无线网络的空间处理	188
8.7.1	Rayleigh 衰落信道下的网络模型	189
8.7.2	数据传输层模型	189
8.7.3	几个基本概念与应用	190
8.8	未来的研究方向	192
	习题	195
	参考文献	195

第 1 章 概 论

网络信息论,作为一个新兴的通信研究领域,已引起国际学术界的广泛重视。为了使读者尽快了解有关网络信息论的发展概况以及未来的研究方向,在本章中,我们简略地介绍这方面的内容,其中包括网络信息论研究的几个基本问题、网络信息论的发展历史、与网络信息论有关的部分应用以及未来的研究方向等。

1.1 网络信息论研究的三个基本问题

近 20 年来,通信技术得到飞速发展,数字通信的概念已被整个社会所接受,电子信息社会化程度越来越高。在世界上,城市信息化与信息网络无缝覆盖已成为当前发展的主流。作为信息交流的载体,信息通信日益成为人们生活中的重要组成部分,像衣、食、住、行一样,被现代人称为第五类生活必需品。随着通信需求程度的日益提高,随之而来的是,对通信信息传输的网络化处理技术的要求也越来越高。因此,信息网络的规划、建设与管理变得越来越重要。

在这样一个大背景下,传统的网络规划、建设与管理以及相应的网络容量的评估方法都难以适应现代通信发展的需求。这是因为传统的通信网络规划主要依据的是 ISO 定义的层次化的网络模型,该网络理论将物理层看成一个信息通信的管道,通过分层屏蔽掉物理层对网络层的影响,以排队理论为基础,通过控制网络流的路由和交换作为信息处理手段,通过评估网络接入接口的容量推算整个网络的承载能力,这样的网络容量估计并非真正体现信息流端到端的网络容量,所以网络的空闲程度以及网络的资源利用效率的评估也是不合理的。另一方面,随着通信业务的种类变化,数据传输逐渐成为通信业务的主要组成部分,话音通信在整个通信业务中的份额越来越小,传统的基于排队理论的业务量分析理论在容量估算中带来的偏差越来越大。因此,基于网络信息流的自相似分析理论开始变得活跃,并得到应用。为此,在理论上,如何准确客观地评估一个现有网络的信息容量成为网络信息论研究的一个重要问题。

考虑传统的话音通信模式,我们知道 Erlang-C 公式^[1]是评估网络交换节点接纳话音业务的重要方法和工具。我们能观察到话音网络容量与信息流模式的关系,以及与业务模式的关系。所谓的容量就是交换节点允许同时接入的电话数目。网络中的信息流是按照开关模式进行的,即电话到达按照 Poisson 分布模式,话音服务的时间间隔按照指数分布进行统计的。在有电话时,整条线路将被一直占用;通话结束后,线路才被释放,转为空闲状态,并且这种空闲状态一直持续到下一个新的通话开始为止。显然,业务对线路的占用模式、业务到达时间分布、业务服务时间分布、交换中心线路的使用策略等都会对整个交换中心的话务容量有影响。由此可见,这种情况下观测到的网络容量实际上是在一个具体的实现模式下的实际利用容量界。它并非是一个话务网络所能承载的总容量^[2~3]。

对现代的数据网络而言,一般情况下,数据业务对线路的占用模式不再是独享线路资源的模式,而是多路数据业务共享线路的模式。一般地讲,可能的共享方式包括时分、频分(波分)、码分以及网络编码等。数据业务的到达时间分布、数据包的服务时间分布、中心节点的交换与转发策略等都与先前讨论的话音业务有很大差别。特别是,在交换节点处有可能出现数据风暴,即单位时间内有大量的数据包同时到达。此外,每条接入交换中心的线路,它们的带宽也可能存在较大的差异。因此,如何将这些因素考虑进去,从中提炼出一个具有一般特征的网络容量的定义成为目前通信学界关注的焦点问题。

当然,网络层在逻辑上所得到的网络容量与利用网络信息论观察所得到的网络容量是有本质差别的。网络信息论讨论的容量问题要远复杂于仅在网络层讨论网络容量的问题,它要将网络层、传输层以及物理链路层多层的特点融合在一起进行讨论,观察网络的整体特征。因此,从信息网络的整体观察一个网络的特征,对网络的运行机制和传输效率的评估无论在理论上还是在应用上都更有价值。实际上,就网络的容量分析而言,从通信网络的整体观察一个网络的容量,也许比从局部观察网络的特征,使问题的解决变得更加容易。

网络信息流容量的估值只是从技术指标上给出了网络的最大承载能力,如何有效地利用网络的承载能力,使其使用效率最大化成为网络信息论研究的第二个重要问题。正如我们研究传统的 Shannon 信息论,对一个给定的点到点的传输信道,如 Gaussian 信道,其信道容量可以由 Shannon 公式得到。此时采用什么样的信息传输模式才能按照信道容量进行传输信息呢?为了回答此问题,Shannon 从信息统计的角度给出了一个结论,对需要传输的信息,只要利用长序列编码器进行编码就能够实现按信道容量进行传输,即信道容量的利用率可以接近百分之百。然而,在通信系统实现时,人们是不可能采用无限长序列的编码模式的。也就是说,所有可实现的信道编码模式都是有限长的。因此,针对有限长的信道编码,人们经过近 50 年的努力,才在 1993 年由 Berrou 等人发现了 Turbo 码(级联卷积编码),此时,信道容量利用率接近百分之百的问题才得到解决。1996 年,LDPC 码(低密度编码)的再发现,使得人们意识到分组码也可以使信道容量利用率接近百分之百。

事实上,网络容量使用效率的最大化问题远比信道容量利用率的最大化实现要复杂得多。其主要原因在于,在网络中信息流的传输除了涉及传输线路的差异之外,还应考虑网络中各个节点对信息流的处理能力以及控制策略。此外,它还涉及网络中信息流承载业务的特征等。因此,如何提高网络的使用效率将成为一个长远而艰巨的研究课题。目前,很多研究主要集中在如何提高网络的传输效率或如何提高网络的吞吐量,而忽略了网络容量的使用效率。这是因为网络容量问题到目前为止仍没有一个明确的解。就一个具体的网络而言,其容量分析也只局限于实际可实现的传输容量。为此,人们直觉地认为提高现有网络的传输效率可以间接地提高网络容量的使用效率。于是,人们在该问题上的研究观点出现了变化。人们更多地开始讨论如何提高现有网络的传输效率,而不再关心网络容量的使用效率。基于这种思想,通信网络的规划和建设不再关心网络容量,而关心能提供的具体通信业务量或服务的客户数目,甚至有时只关心是否能满足客户的传输业务要求。从这个层面上讲,通信资源的使用效率在具体的工程应用中是较低的,这与通信网络的规划和建设最初设定的目标有直接的联系。因此,对一些具体的通信网络而言,我们在关心提高网络的吞吐量或提高现有网络使用效率的同时,应利用一些网络信息论提供的网络容量界评估网络容量的使用效率,并以此为根据,适当调整网络中信息流的传输控制策略,进一步提高网络资源

的使用效率。

除了网络的容量和容量使用效率问题之外,网络信息论关心的问题是通信网络的工程使用问题。因为目前通信网络的种类繁多,每一类网络都是特定历史阶段发展的产物,也都有特定的设计理念和特定的用途,所以,不同网络之间的互联互通,以及网络资源互补等研究也成为网络信息化处理的重要研究课题。它与前两个问题的不同之处在于,前两个问题可以从逻辑上自然得到,而它是网络信息化发展过程中派生出的新问题,也是现代网络信息论在工程实现和理论研究上的第三个重要问题,或许也是最重要、研究难度最大的问题。实际上,早在 20 世纪 80 年代,人们就提出了综合业务数字网的概念,之后又开始了讨论电话网、数据网以及广播网三网合一的问题,直到今天讨论的基于软交换的下一代网络技术等。这种网络整合和归一化处理,即信息传输体系一体化,一直是人们追求的目标。近年,由于 IP 技术的快速发展,人们也开始期望基于全 IP 的通信网络建设。然而由于 IP 技术本身的一些缺陷,如网络管理方面的不可控性,业务服务质量保障方面的不确定性等,改进 IP 技术的研究也成为目前通信学界研究的主流方向之一。此外,由于电子器件的革命性进展,复杂的数据存储、交换、计算等处理能力得到飞速提高,使得通信组网方式变得更加灵活,特别是在无线通信方面。于是,大量利用无线通信方式进行组网的特殊网络类型不断地出现,包括无线局域网、无线城域网、无线个域网、无线自组织网络以及无线传感器网络等。这些特定的网络,有时可以自成一体,有时它们作为一个接入网需要与骨干的通信网络进行连接完成信息通信。显然,特定的通信网络以及网络互连的容量分析和可实现的最大通信效率问题也是网络信息论关心的重要问题。

1.2 网络信息论的发展历史

网络信息论,作为一个新兴的通信研究领域,越来越受到国际学术界的重视。在美国和欧洲,很多著名的大学和研究机构都在开展这方面的研究。特别值得注意的是,自数字通信的概念和 Shannon 信息论诞生之后不久,信息传输网络化就引起学术界的注意。在 1956 年,Shannon 和他的同事就以“关于通信网络最大流的注记”为题开始讨论网络信息传输理论。由于网络容量分析的难度以及当时通信网络主要是通过有线传输话音,因此,这一研究并没有引起学术界的广泛关注。而通信网络的容量分析问题变为单交换节点的话务量分析问题,于是出现了将物理层完全屏蔽的排队论模型,它只关心网络中话务的到达时间分布以及网络中话务服务时间的分布等参数,这就是所谓的 Erlang 公式。然而,将物理层纳入网络容量的分析范畴的研究并没有停止,只是受学术界的关注程度变弱了,从事这方面研究的学者人数较少。在 20 世纪 60 年代至 70 年代,一些学者,如 Cover, Liao, Wyner 和 Gallager 相继讨论了多址信道、广播信道、中继信道和干扰信道以及接力信道等容量问题^[4]。这些讨论主要是沿用了传统的 Shannon 信道容量的讨论方法,因此,所采用的具体分析策略以及得到的结论都难以推广到一般性的网络模型中。关于网络信息容量分析的研究虽然一直没有停止,但是其进展非常缓慢,这种状态一直持续到 20 世纪末。在刚步入 21 世纪的最初几年内,网络信息论的研究取得了突破性的进展,其代表作和主要贡献包括香港中文大学杨伟豪教授和李硕彦教授等在 IEEE 信息论会刊上发表的“网络信息流”(network information flow)^[5]和“线性网络编码”(linear network coding)^[6];美国学者 Gupta 和 Kumar 教授提出

的无线网络的传输容量 (transport capacity)^[7]; Xie 和 Kumar 提出的 scaling law and optimal operation^[8]; Grossglauser 和 Tse 给出的移动性可以增加无线自组织网络容量^[9]等, 这些成果的出现拓宽了网络信息论的研究方法和研究视角。概括地讲, 这些研究工作的主要贡献可以归结为如下几点:

- (1) 网络编码将编码处理的功能引入网络层, 拓宽了传统的网络层接收转发功能。
- (2) 网络层的编码处理可以用于提高组播网络的可实现传输速率, 使之达到或接近网络的最大流量界。
- (3) 网络层编码在网络管理以及抗链路差错方面, 提供了信息冗余度, 这使得物理层的信道编码理论得到扩展和推广。
- (4) 线性网络编码的出现意味着线性空间处理技术可以满足网络编码在组播容量和网络链路管理上的基本需求。
- (5) 传输容量概念的出现将无线通信中信息传输速率与传输距离考虑在一起, 拓宽了信息容量的定义, 为无线通信的研究提供了新的度量标准。
- (6) 网络的缩放规则有助于揭示网络规模变化与网络性能特征的内在联系, 它已经成为讨论无线自组织网络的重要方法之一。
- (7) “移动性可以增加无线自组织网络的容量”表明空间资源的利用可以进一步提升网络的传输容量界。

这些工作的贡献也许还有更多的内在含义, 我们希望上述总结能起到抛砖引玉的作用, 使读者能联想到更多的理论意义和使用价值。在这些工作成果的驱动下, 目前有很多学者开始关注并从事相关的研究, 特别是在网络编码、无线自组织网络、无线传感器网络等方面。我国一些大学和研究所也投入到这些课题的研究中, 并得到国家自然科学基金和国家 863 等项目的资助。可以预见, 在不久的将来, 这些方面的研究一定会取得更加丰硕的成果。

1.3 网络信息论的应用

网络信息论的研究虽然有很久的历史, 但仍缺乏像传统的 Shannon 信息论那样具有完整的理论体系框架, 因此, 这些研究留给人们的印象就是理论体系不完整, 理论结果都属于纸上谈兵, 没有明确的应用背景和使用价值。尽管如此, 网络信息论的部分结果却被自觉或不自觉地应用着, 并且取得了明显的性能增益和社会效应。下面就几个具体的例子进行说明。

例 1.1 网络信息下载

利用计算机网络进行信息搜索和信息下载已成为城市居民信息获取的主要手段之一, 并被人们广泛接受。在进行信息搜索时, 人们自然会联想到 google, sohu, Yahoo, baidu 等网络搜索工具。利用这些搜索工具, 只需输入一些关键词就可查找到相关的网页或具体的有关文本内容。在多数情况下, 人们会对搜索到的文件或网页进行下载和保存。此时, 如果利用传统的信息下载工具, 即所谓的基于纯路由的点到点下载工具, 文件的下载速度有时会非常慢, 尤其是在一些网络使用的高峰时段。这是因为此时有很多用户同时在登录或访问某一网站。用户们相互竞争网络资源, 抢占传输信道势必造成网络拥塞。近期 Microsoft 开发了一种新的网络信息下载软件 Avalanche, 又称为“雪崩”软件。利用该软件, 在 P2P 模

式下可以成倍地提高信息下载的速率。特别地,对某一共同感兴趣的信息文件,网络中使用该软件下载信息的人数越多,信息下载的平均速率也越高。实现这一网络信息下载功能的软件所采用的基本原理就是网络编码。因为用户对各自得到的信息进行编码,然后以编码模式转发给其他用户,这样可以保证每个用户不需要直接从被访问的网站下载所有的相关数据,只需下载其中的一部分即可。至于其他的数据,用户可以与其他用户相互交换编码数据,通过解码得到。从网络传送数据的角度可知,传统的点到点的纯路由下载模式要求每个用户都建立自己的访问路由,这样就会导致用户相互竞争访问资源,造成信息下载速率下降。尽管从信息传输的角度看,多用户共享一个信息源的传输是一种网络组播模式,但在本质上它是通过时分方式利用多个单播组合实现的。与此相反的是,“雪崩”软件利用了网络编码的手段,用户间通过交换编码数据,这样无形中建立了部分联系,也就是说,建立了相关的路由。从原理上讲,它除了利用点到点的纯路由策略外,还利用了网络中其他的路径相互交换信息,从而提升了每个用户平均信息下载速率。除此之外,这种下载软件的使用,也间接地消除了部分网络链路故障或部分链路拥塞甚至部分用户自动下线造成的影响,提升了网络界面的友好程度。

从网络信息论的角度看待这一应用,我们可以知道,对组播业务,当采用网络编码时,其网络传输容量应等于被访问网站和多个用户之间的最大流,该数值一般情况下要远大于被访问网站与每个用户之间建立的单路由路径的传输速率,因此,采用“雪崩”软件的信息下载速率要远高于利用传统的基于纯路由的点到点的信息下载速率。

图 1.1 给出了一个网络信息下载拓扑图,其中标号节点为最终的信息下载节点。显然,利用网络编码,网络中除信源服务器之外的每一个节点都可以对自己接收到的信息进行编码(组合),然后再将编码后的信息发送到它的输出链路上。对每个节点而言,不同的输出链路可以采用不同的编码,也可以采用相同的编码。例如,目的地节点 3 可以通过两条不同的路径下载信息,因此它有三种不同的选择,一种是在这两条路径中选择一条下载速度快的路径,另一种是同时利用两条路径下载信息,不同的路径下载不同的部分,还有一种选择就是利用网络编码,防止中间节点窃听部分下载信息(具体编码方法,请参看第 6 章网络编码部分)。对节点 4,它存在多条下载路径,但不相交路径只有三条,因此,它最大的下载速率只能同时利用三条不相交的路径。对节点 1 和节点 2 而言,它们有一部分下载路径要通过节点 4,在信息下载过程中,如果节点 4 在自己下载完成后就退出网

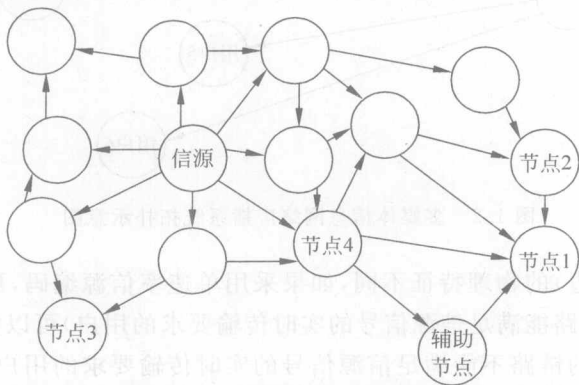


图 1.1 一个网络信息下载拓扑图

络,不再帮助节点 1 和节点 2 下载信息,此时可以看出,如果采用网络编码,那么节点 1 和节点 2 仍能通过其他路径下载自己需要的剩余部分信息,并利用节点 4 离开前的转移到黑色节点(辅助节点)的部分信息帮助完成信息下载。

例 1.2 多媒体信息网络广播系统

信息广播是一种重要的应用,提到传统的信息广播系统,人们自然就会联想到无线电广播和电视广播。然而,近年,通过计算机网络进行多媒体信息网络广播已成为一种非常重要的应用。这种通过计算机网络进行多媒体信息广播之所以流行,其主要原因在于,现代的计算机网络已变得越来越普及。可以通过有线接入或无线接入将个人计算机与网络进行连接。例如,某些公司职员要利用午餐休息时间观看网络电影、体育比赛或新闻报道等,他们可以利用个人计算机通过登入相关网站收看。此时,传送的信息一般情况下都是多媒体信息,包括语音、文字和图像,甚至包括一些数据等。对实时的信息传输,由于每个用户所处的地理位置不同,他们与该多媒体信息播送网站之间建立的传输链路也不同,而每条链路的传输带宽以及拥塞程度也存在差异,因此,这些用户所利用的网络路径难于保证它们都能实时地收到满意的信号。如图 1.2 所示,链路线条的长短表示链路最大可行的信息传输速率的大小,其中较短线条表示较高传输速度。显然,对那些建立的网络路径带宽较小的用户或网络路径拥塞程度较高的用户而言,如果不采用特殊的信息处理手段,一方面,他们很难收看到该多媒体网站播放的信息,如图 1.2 中用户 3 和用户 4。另一方面,这些用户在下载信息时,如果下载过程出现链路中断,那么下载到的部分信息也需要丢掉,这是由于只利用部分信息是无法正确解码的。这样势必造成网络资源的浪费,同时也影响了其他网络用户对网络资源的使用。为了避免网络资源的浪费,也能使那些具有链路性能较差的网络用户能得到较满意的信息收看效果,需要建立适应网络不同传输路径特征的信源编码,这也就是多层信源编码或多速率信源编码产生的主要原因之一。

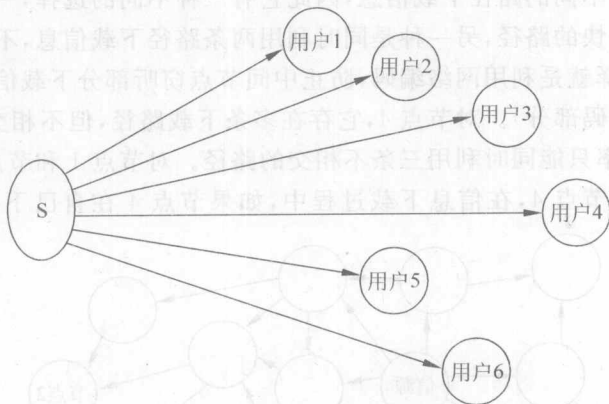


图 1.2 多媒体信息网络广播系统拓扑示意图

由于每条链路(边)的物理特征不同,如果采用单速率信源编码,那些拥有链路传输性能好的用户(建立的链路能满足信源信号的实时传输要求的用户)可以实时地访问该多媒体网站;而对那些建立的链路不能满足信源信号的实时传输要求的用户,他们无法实时地接收该多媒体网站广播的信息。如果采用多层信源编码,使信息的传输依据网络链路的特征

实时调整,将信息按其重要程度进行分层编码并规定适当的传输优先级,此时,那些建立起的链路能满足整个信源信号实时传输要求的用户可以实时地访问该多媒体网站,接收全部该多媒体网站广播的信息;而那些建立起的链路不能满足整个信源信号实时传输要求的用户,可以选择与自己传输链路匹配的信源发送信息,接收最重要的部分信息,从而也能得到较满意的信息接收效果。

事实上,这一问题可以从网络信息论的角度进行解释,即对广播网络,网络拓扑可以抽象为一个星形拓扑。也就是说,多媒体网站作为中心节点,每个用户作为一个边缘节点,在多媒体网站和每个用户之间建立的每条链路看作是星形拓扑的一条边。为了能适应网络的特征,使得网络广播的总信息速率达到最大,需要建立适合每条链路的信息传输模式,即采用多速率或多层信源信息编码,使得不同的链路传输不同的编码信息。显然,由于这些链路特征的复杂性,不可能使每条链路都具有与之匹配的信源编码,因此,工程应用中利用的多速率或多层信源信息编码只是一种可以逼近星形网络传输容量的实现手段。目前,这种概念也被应用到无线信息广播中,相应的分层信源编码转化为不同分辨率的信源或信道编码。

例 1.3 基于传感器的信息收集系统

通常情况下,可以利用传感器进行信息收集。例如,在海洋中利用传感器观察鱼群的活动规律;在农业生产中,利用传感器观察农田的病虫害情况;在林业生产中,利用传感器观察森林的病虫害情况、进行森林火灾预报等。传感器信息收集系统也可以抽象为一个星形网络,每个具体的传感器可以看作一个网络的边缘节点,它们负责将各自收集到的信息转发给收集节点;而收集节点可以看作网络的中心节点,主要负责将各个传感器节点收集到的信息经过处理转发到观测中心。传感器节点与收集中心节点的传输线路看作网络拓扑中的边。

如果只考虑传感器信息收集网络的实现,显然只需要建立一个简单的传输协议,并用于负责节点间的信息传输即可。然而,这样的物理实现,势必会造成网络资源的大量浪费。从信息收集的角度讲,由于传感器之间相对物理位置的关系,某些传感器观察到的信息可能具有很强的相关性,例如,图 1.3 中传感器 1 和传感器 2 的物理位置很近,它们收集到的(地理)信息可能具有很大的相关性。因此,它们传送给收集中心的数据信息也具有强相关性,收集中心收到的数据中含有大量的冗余信息,这样会造成网络资源的大量浪费,也不利于网

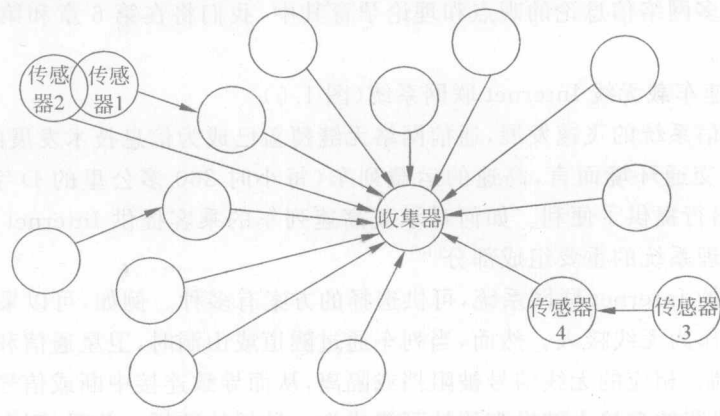


图 1.3 传感器网络示意图

络传输协议的优化。另一方面,也会造成传感器发送功率的浪费,特别是对采用电池供电的传感器,其使用寿命将大大降低,进一步导致传感器网络使用寿命的下降。此外,信息收集中心也有可能将这些具有强相关性的信息发送给观察中心,增加了信息观察中心处理数据的复杂度。如果采用网络信息论的观点,人们自然就会从信息网络的角度考虑信息收集的传输与处理。在传感器信息收集时,应考虑传感器之间观察数据的相关性,通过传感器之间的信息交流,去除大量的观察数据之间的相关性,然后将弱相关的观察数据发送给信息收集中心。一方面,由于传感器发送数据量的减少,传感器的平均发送功率会降低。从统计的角度讲,对采用电池供电的传感器,这样可以延长整个传感器网络的使用寿命。另一方面,由于网络中需要传输的数据量的下降,网络传输协议的设计可以适当地简化,并有可能得到进一步的优化,使得网络的信息传输效率得到提高。对数据处理中心而言,数据处理的复杂度也可以降低,使实时数据处理成为可能。

例 1.4 分布式无线传输系统

在传统的网络信息论中,基于三节点的无线接力传输模型,如图 1.4 所示,受到通信学界的广泛关注。从这一研究中得到的重要启示在于,如果在通信的信源节点和接力节点之间建立了节点协作关系,那么,与直接传输相比,可以提高信源节点的信息传输速率。从网络信息论的角度看,直接传输的信息速率最大值为信源节点与目的地之间的信道容量。而采用接力节点辅助传输,则信息传输的容量等于信源节点与目的地之间的信道容量加上一个附加量。这个附加量就是信源节点到接力节点之间的信道容量与接力节点到目的地节点之间的信道容量中的较小值。显然,与直接传输相比,接力辅助传输带来了使用效率的增加。该增加量就是我们通常所说的信息传输的多址分集增益。

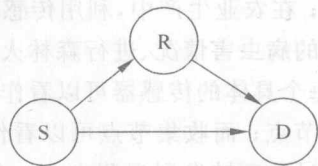


图 1.4 三节点的无线接力传输模型图

S: 信源节点; R: 中继接力节点;
D: 目的地节点

基于上述原理,在分布式无线传输系统中,也引入了节点协作关系,节点之间可以通过某种协议,相互之间互为接力节点负责转发邻居节点提供的信息,从而提高信源节点的信息传输速率。此处,只简单地介绍了网络信息论的一个应用观点。可以预见,在分布式无线传输系统中仍有很多网络信息论的观点和理论孕育其中,我们将在第 6 章和第 8 章结合具体的应用予以介绍。

例 1.5 高速车载无线 Internet 联网系统(图 1.5)

随着无线通信系统的飞速发展,通信网络无缝覆盖已成为信息技术发展的必然要求,正处于实施中。从交通环境而言,高速的运载列车(每小时 260 多公里的 D 字头列车)的出现,为人们旅游出行提供了便利。如何对承载高速列车的乘客提供 Internet 联网服务成为开发新型列车管理系统的重要组成部分^[10]。

对高速车载的 Internet 联网系统,可供选择的方案有多种。例如,可以采用卫星通信或无线蜂窝通信等作为无线接入。然而,当列车通过隧道或山洞时,卫星通信和无线蜂窝通信将受到严重的限制,相应的无线信号被阻挡或隔离,从而导致连接中断或信号严重衰落。为此,沿列车运行线路的多接力转发直放站可能成为一种好的选择。此时,列车上安装的天线系统与地面沿路天线系统之间就形成了一个无线传输网络。由于列车的运行速度较高,列