

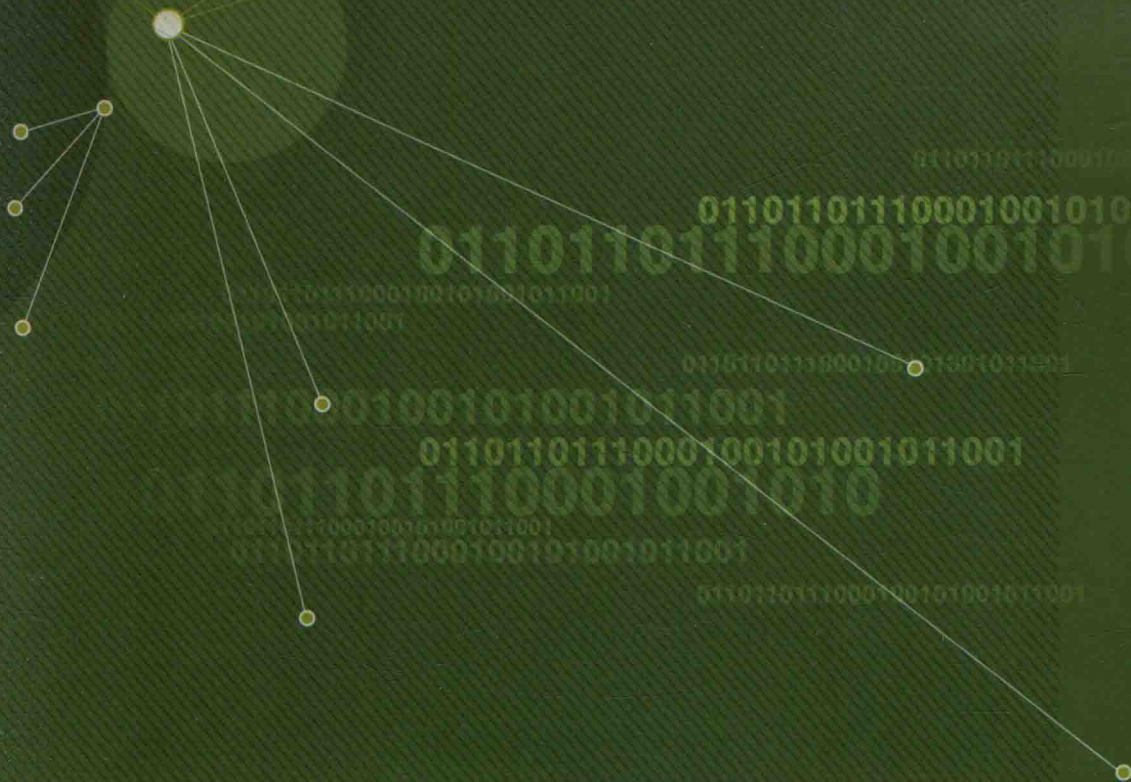


出版社出版基金资助项目
CHUBANSHECHUBANJINZHUXIANGMU

复杂网络传输容量 分析与优化

FUZA WANGLUO CHUANSHU RONGLIANG FENXI YU YOUHUA

◎ 蒋忠元 梁满贵 著



北京交通大学出版社
<http://www.bjtu.cn>

复杂网络传输容量 分析与优化

FUZA WANGLUO CHUANSHU RONGLIANG FENXI YU YOUHUA

蒋忠元 梁满贵 著

北京交通大学出版社

· 北京 ·

内 容 简 介

自从小世界与无标度网络模型问世以来, 复杂网络研究进入了快速发展的阶段, 研究成果不断涌现。复杂网络研究存在于许多学科当中, 数学、生物学、工程与科学技术基础等诸多学科领域内的研究者们已经对复杂网络展开了一系列的研究, 计算机科学技术学科也不例外。该学科领域内的研究者们主要针对计算机复杂网络中存在的问题, 比如复杂网络拥塞展开了研究。本书重点剖析了优化网络传输容量的瓶颈, 从复杂网络角度定义了网络传输容量, 重点分析了导致复杂网络发生拥塞的原因, 分别从复杂网络路由策略、复杂网络拓扑结构、复杂网络带宽分配三个方面提出了优化策略, 以提升复杂网络承载数据的能力和传输效率。此外, 本书还进一步探讨了优化策略在实际网络中的可行性。本书的相关研究成果可为当前网络优化或下一代网络规划提供参考。

本书可作为计算机科学与技术专业研究生或相关领域科研工作人员的参考书籍。

版权所有, 侵权必究。

图书在版编目 (CIP) 数据

复杂网络传输容量分析与优化 / 蒋忠元, 梁满贵著. — 北京: 北京交通大学出版社, 2016. 3

ISBN 978-7-5121-2474-5

I. ① 复… II. ① 蒋… ② 梁… III. ① 计算机网络-研究
IV. ① TP393

中国版本图书馆 CIP 数据核字 (2015) 第 270619 号

复杂网络传输容量分析与优化

FUZA WANGLUO CHUANSHU RONGLIANG FENXI YU YOUHUA

责任编辑: 韩 乐 助理编辑: 严慧明

出版发行: 北京交通大学出版社

电话: 010-51686414 <http://www.bjtu.cn>

地 址: 北京市海淀区高梁桥斜街 44 号 邮编: 100044

印 刷 者: 北京艺堂印刷有限公司

经 销: 全国新华书店

开 本: 170 mm×235 mm 印张: 9 字数: 126 千字

版 次: 2016 年 3 月第 1 版 2016 年 3 月第 1 次印刷

书 号: ISBN 978-7-5121-2474-5/TP·828

定 价: 36.00 元

本书如有质量问题, 请向北京交通大学出版社质监组反映。对您的意见和批评, 我们表示欢迎和感谢。
投诉电话: 010-51686043, 51686008; 传真: 010-62225406; E-mail: press@bjtu.edu.cn。

前言

在当今信息时代，人们的生活与形形色色的网络密切相关，无论是工作、学习还是日常生活，都已深深依赖功能各异的网络。通信网络为信息传输、资源共享提供基础设施。万维网（WWW）由互相链接的超文本文件组成，为用户提供丰富的信息资源。人们通过 QQ、微信、微博、MSN、人人网、Facebook、Twitter、Google+、Myspace、LinkedIn、VKontakte 等社交网络来认识新朋友，分享好友趣事，关注好友状态等。导航系统依赖国家公路网规划，为用户提供便利、准确的导航信息。民航飞机依赖航线网络进行调度与飞行规划。还有生态网络、蛋白质网络、新陈代谢网络、神经网络等都可以通过复杂网络模型进行建模，并采用复杂网络理论对实际网络中存在的问题展开研究，比如揭示交通网络或通信网络的网络拥塞原理并提出减缓网络拥塞的方法，探究局部故障如何导致电网大面积停电事故，研究病毒在网络中的传播机理及遏制机制等。

本书以计算机通信网络为背景，以研究计算机通信网络中实际存在的网络拥塞问题为出发点，通过复杂网络理论分析问题、描述问题，对问题进行建模，探寻解决方法。首先系统分析导致网络拥塞的原因，其次综合分析已有的减缓网络拥塞的方法，最后重点介绍作者对该问题的研究贡献。全书共分为 5 章：第 1 章介绍本书的研究背景，分析导致网络拥塞的三个方面的——路由策略、网络拓扑结构、资源分布，并介绍相关网络模型及评估参数，系统综述复杂网络的增容途径；

第2章介绍作者在复杂网络路由策略优化方面的工作，共研究了五种路由优化策略；第3章从优化网络拓扑结构角度出发介绍作者的研究内容，主要包括增边策略研究、重连边策略研究、两层网络模型上的删边机制研究；第4章主要研究在网络链路带宽资源总量一定时，如何合理地为每条链路分配带宽，以达到提高网络传输效率的目的，为未来网络或下一代网络规划提供依据；第5章对整个研究内容进行了总结并提出了对未来工作的展望。

本书由西安电子科技大学蒋忠元老师与北京交通大学梁满贵教授共同完成，其特点主要体现在以下几个方面。

1. 本书通过剖析导致网络拥塞的原因来探索减缓网络拥塞的策略，并以理论分析与实验仿真分析为主线，便于读者梳理思路。

2. 本书第1章抛出研究主题，剖析问题原因，整体介绍可能解决问题的方法，后续第2、3、4章依次针对导致网络拥塞的三个方面提出优化机制，第5章对全书作了总结。每章节安排以树形结构呈现，力求层次清晰，便于读者阅读与理解。

3. 本书总结了复杂网络领域的学者对优化网络传输容量这一问题的研究进展，并重点介绍作者对该问题的研究贡献，其目的在于与同行学者进行交流，希望能为本研究问题的发展略尽绵薄之力。

由于作者能力和水平有限，书中不当之处或错误在所难免，敬请专家学者及广大读者批评指正。

作者

2016年2月



目 录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 主要模型及参数	3
1.2.1 网络模型	3
1.2.2 流量模型	5
1.2.3 主要参数	7
1.3 网络增容途径	11
1.3.1 路由策略	11
1.3.2 网络拓扑结构	15
1.3.3 资源分配	18
1.4 本书主要研究内容及结构	21
第 2 章 复杂网络路由策略优化	24
2.1 基于介数的有效路由策略	24
2.1.1 路由算法描述	25
2.1.2 实验仿真与分析	27
2.2 基于链路拥塞的加权路由策略	33
2.2.1 路由算法描述	33
2.2.2 实验仿真与分析	34

2.3	增量路由策略	38
2.3.1	路由算法描述	39
2.3.2	实验仿真与分析	41
2.4	基于向量交换的全局动态路由策略	47
2.4.1	路由算法描述	50
2.4.2	实验仿真与分析	50
2.4.3	扩展与应用	55
2.5	基于两种流量级别的动态源路由策略	58
2.5.1	模型与路由算法描述	59
2.5.2	实验仿真与分析	60
2.6	本章小结	64
第 3 章	复杂网络拓扑结构优化	67
3.1	增边策略研究	67
3.1.1	增边策略	68
3.1.2	实验仿真与分析	71
3.2	重连边策略研究	82
3.2.1	重连边策略	82
3.2.2	实验仿真与分析	83
3.2.3	扩展研究	88
3.3	两层网络模型上的删边机制研究	90
3.3.1	两层网络模型与删边机制	90
3.3.2	实验仿真与分析	92

3.4 本章小结	95
第 4 章 复杂网络带宽分配优化	96
4.1 基于介数的带宽分配策略	96
4.1.1 带宽分配策略	97
4.1.2 实验仿真与分析	97
4.2 全局动态带宽分配策略	101
4.2.1 带宽分配策略	101
4.2.2 实验仿真与分析	102
4.3 本章小结	107
第 5 章 研究总结与展望	108
5.1 研究总结	108
5.2 研究展望	110
附录 A 基于节点拥塞流量模型仿真实验介绍	112
附录 B 基于链路拥塞流量模型仿真实验介绍	116
参考文献	120



第 1 章

绪 论

1.1 研究背景及意义

在当今信息时代，网络无处不在，我们的日常生活与网络息息相关，比如因特网（路由器为节点，物理链路为边）、万维网（超文件为节点，超文件之间的超链接为边）、高速公路网（高速公路枢纽为节点，高速公路为边）、飞机航线网（城市为节点，城市之间的航线为边）、人际关系网（人为节点，人与人之间的联系为边）等。一直以来，网络研究都是计算机通信、交通运输等领域的研究热点。网络最重要的功能是用来传输其所能承载的对象，比如因特网中的数据包、城市交通网中的交通车辆等。复杂网络中通常用网络传输容量（traffic capacity）来衡量网络的传输能力，即传输容量越大对数据传输越有利。以因特网为例，随着用户数量的日益增长及数据量呈指数级的增长，网络被要求有更高的数据承载能力。然而，在实际网络中经常因数据量过载而发生拥塞，给人们带来许多不便与经济损失。近年来，基于已有的网络结构来研究如何以较低的代价提高或优化网络传输容量成为国内外的研究热点。网络的传输动力学研究及网络传输容量的分析与

优化是减少网络拥塞的必要手段。

网络拥塞一般最先发生在网络中的一些中心(hub)节点上,比如因特网中的路由器^[1-2]或者高速公路网中的高速公路枢纽,之后网络拥塞进一步扩大,最终导致网络整体拥塞甚至造成网络瘫痪。已有不少研究人员开始研究如何控制通信网络中的拥塞^[3-18],但是这些研究一般是在随机构造的网络中进行,而不考虑实际的网络结构。从另外一个角度来说,网络的完整拓扑结构很难获取得到,因而缺少大规模的实际网络拓扑结构,这给全面研究网络动态性能带来一定的困难。

自从1998年Watts和Strogatz在*Nature*杂志上发表文章揭示了复杂网络的小世界特性(small-world)^[19]与1999年Barabási和Albert在*Science*杂志上发表文章揭示了复杂网络的无标度特性(scale-free)^[20-21]后,大量研究证实很多实际网络都具有小世界特性与无标度特性,比如因特网、高速公路网、万维网、社交网络等。从此,采用复杂网络模型对实际网络结构进行建模与研究逐渐成为热点^[22-54],比如交通网^[24]、细胞网络^[26-28]、新陈代谢网络^[34]、因特网^[44-54]等。利用复杂网络理论与研究方法研究实际网络的动态特性也逐渐成为新的研究热点。

复杂网络传输容量的优化是复杂网络最基本也是最重要的问题之一,近年来关于优化复杂网络传输容量的研究方兴未艾。定义网络传输容量的模型被称为流量模型,包括用户端数据生成、传输过程、网络资源的调度使用等。其中,网络中的数据生成于用户端也终止于用户端,用户是数据的主导者,因此用户端的行为也极大地影响网络传输容量;传输过程由路由算法来选定路径;网络资源的调度使用包括队列调度、节点容量使用、带宽使用等。因此,网络传输容量是综合考虑路由算法、网络拓扑结构、资源分布的共同作用来评估网络性能的参数。最短路径路由算法在实际网络中被广泛应用^[55-56]。假设网络中的所有节点都对等,即所有节点同时扮演诸如计算机网络中的主机和路由器角色,而且各个节点所具有的资源也相同。由于实际网络拓扑结构具有无标度特性,因此在最短路径路由算法下,网络的中心节点极易发生拥塞,最终拥塞扩散至整个网络。不少学

者提出优化网络路由策略, 根据实际网络拓扑结构, 设计有效的路由策略使网络中的流量尽量均匀地分散到各个节点上^[57]; 也有不少学者认为实际网络拓扑结构的不均匀性导致网络的整体性能下降, 因而提出通过调整网络拓扑结构的方法来提高网络传输容量^[57-58]。然而, 上述优化路由策略或者网络拓扑结构的研究方法容易忽略网络中的资源分配和分布问题, 比如节点处理能力、链路带宽、缓存大小等资源的合理分配。其实, 实际网络中的资源非常有限而且分布不均匀。因此, 可以根据网络拓扑结构与所使用的路由算法进行合理的资源分配, 使网络的中心节点或者负载过大的链路分配到较多的资源, 以保证网络的整体资源被充分而有效地利用。

研究提高网络传输容量的策略具有非常重要的现实意义。首先, 通过建立合理模型对大规模实际网络进行合理评估, 为全面理解和分析网络的动态性能奠定基础; 其次, 利用现有的网络资源, 无论是优化路由策略、调整网络拓扑结构, 还是合理地进行资源再分配, 都能从宏观上大幅地提高网络传输容量, 极大地缓解网络拥塞, 提高网络承载能力; 第三, 以较低代价优化网络传输容量可以显著地改善用户上网体验, 并且可以给社会带来极大的经济效益。

1.2 主要模型及参数

1.2.1 网络模型

1.2.1.1 BA 无标度网络模型

本书主要采用最为广泛使用的 Barabási-Albert 无标度网络模型^[20-21] (以下简称 BA 网络模型) 对实际网络进行建模。BA 网络模型考虑到了实际网络的两个重要特性, 即生成 BA 网络的规则有如下两点。

(1) 增长 (growth) 特性: 从一个具有 m_0 个节点的网络开始, 每次新增一个节点, 并且连接到 m 个已经存在的节点上 ($m \leq m_0$)。增长特性保证网络的规模在不断地扩大。

(2) 优先连接 (preferential attachment) 特性: 新增节点与一个已经存在的节点 i 相连接的概率 Π_i 与节点 i 的度 k_i 、节点 j 的度 k_j 之间满足如下关系:

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

式中, 变量 j 取遍网络中已存在的所有老节点。

新节点倾向于与那些节点连接度大的老节点相连, 这种现象被称为“富者更富” (rich get richer) 或“马太效应” (matthew effect)。目前对 BA 网络的度分布的理论研究方法主要有连续场理论 (continuum theory)^[20]、主方程法^[59]和速率方程法^[60], 其中主方程法与速率方程法等价, 得到的 BA 网络的度分布函数为:

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \propto 2m^2 k^{-3} \quad (1-2)$$

式中, 度分布 $P(k)$ 表示任意一个节点的度大小为 k 的概率, 详细定义见 1.2.3.1。

式 1-2 表明 BA 网络的度分布函数可以由幂指数为 3 的幂律函数近似描述, 如图 1-1 所示。

1.2.1.2 ER 随机网络模型

1960 年, 在小世界与 BA 网络模型问世之前, Erdős 和 Rényi 就已经提出了 ER 随机网络模型^[61] (以下简称为 ER 网络模型)。ER 网络模型的构建过程为: 在具有 N 个孤立节点的网络中, 任意两个节点之间以一定的概率 p 连接一条边组成网络, 一般用 $G(N, p)$ 来表示。ER 网络模型是一种典型的复杂网络模型, 被广泛应用于学习和研究之中, 其度分布 $p(k)$ 服从 Poisson 分布, 如图 1-2 所示。其中选用网络的平均度用 $\langle k \rangle$ 来表示。

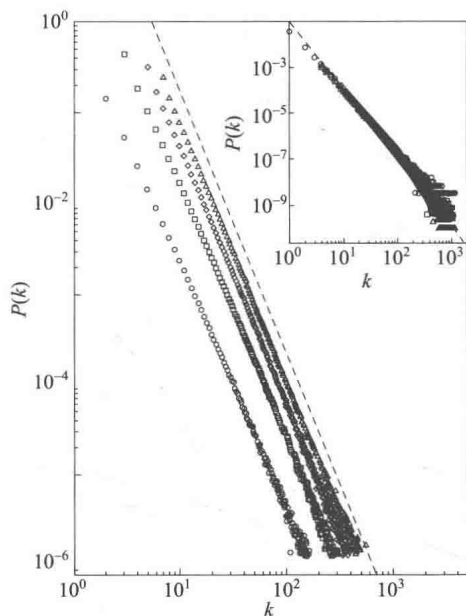


图 1-1 BA 网络中节点度大小为 k 的概率分布图^[21]

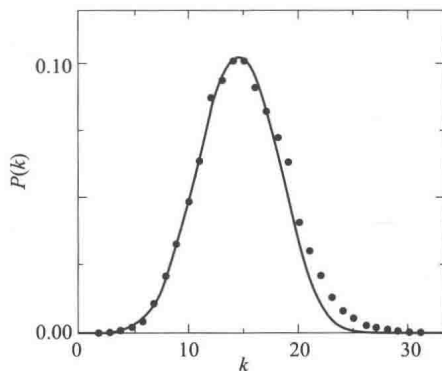


图 1-2 ER 网络中节点度大小为 k 的概率分布图^[22]

注：选用网络的网络规模 $N = 1\,000$ ，平均度 $\langle k \rangle = 4$ 。

从图 1-2 可以看出，ER 网络是一种同质网络，节点度在平均度附近的节点数最多，与图 1-1 中 BA 异质网络结构形成较明显的对比。

本书的研究与仿真实验将采用 BA 网络与 ER 网络两种网络模型。

1.2.2 流量模型

复杂网络的传输容量，一般也被称为网络容量或者网络吞吐量，是衡量网络传输数据能力的重要指标。衡量网络传输容量的模型被称为流量模型，本书将采用复杂网络研究中最常用的两种流量模型——基于节点拥塞的流量模型和基于链路拥塞的流量模型^[62-65]。对于每个给定的复杂网络，假设每个节点同时扮演网络中的路由器角色与主机角色，即每个节点都具有产生、接收、转发数据包的

能力。整个网络每个时刻随机产生 R 个数据包（数据包生成率为 R ），这 R 个数据包的源地址和目的地址都随机选取，并且同一个数据包的源地址与目的地址不同。每个节点处理数据包的能力为 C ，即每个节点每个时刻最多可以处理 C 个数据包。每个节点的缓存队列长度可以无限大，并且数据包到达节点后在每个队列上都采用先进先出（first-in-first-out, FIFO）的排队规则。每个数据包根据一定的路由策略被转发到下一跳节点上。如果下一跳节点不是目的节点，数据包将被保存到下游节点的缓存队列中。数据包一旦到达目的节点，将被立即从网络中删除。

在给定网络拓扑结构、路由策略与资源分布的情况下，不同的 R 值给网络带来的流量负载及拥塞情况也不同，一般我们关注的是数据包生成率 R 的一个关键值 R_c 。当 $R < R_c$ 时，网络处于无拥塞的稳定状态，此时每个时刻进入网络的数据包量与流出网络（被删除）的数据包量持平；当 $R > R_c$ 时，网络中的总数据包量呈现无限增长的趋势，网络中的部分节点开始出现拥塞，即每个时刻流入网络的总数据包量大于流出的总数据包量，流入网络中的数据包不能被及时地送达。一般使用有序参数（order parameter）来刻画网络的状态转换^[6]：

$$H(R) = \lim_{t \rightarrow \infty} \frac{C < \Delta W >}{R \Delta t} \quad (1-3)$$

式中： $\Delta W = W(t + \Delta t) - W(t)$ ， $W(t)$ 表示 t 时刻网络中的总数据包量， ΔW 则表示在 Δt 时间内网络中总数据包量的增长量。有序参数 $H(R)$ 表示数据包生成率为 R 时，网络中每个时刻数据包总量的增长速率，有时也用 η 来表示。数据包生成率 R 有时用 ρ 来表示。当 $R < R_c$ 时， $H(R) \approx 0$ ，网络处于无拥塞状态；当 $R > R_c$ 时， $H(R) > 0$ ，网络中的数据包总量不断增加，网络处于拥塞状态。所以 R_c 为使网络处于无拥塞状态时最大的 R 值，即网络传输容量。

上述网络模型中所有数据包的源地址与目的地址都是随机选取的，所有节点都被等概率地选取为源节点或目的节点。不少研究人员认为，在许多实际网络中，

节点产生数据包的能力并不是完全相同的,或者数据包所到达的目的节点的选取并不是等概率的,而是与节点的度等因素相关。他们认为节点度越大,产生数据包的能力越强,或者被选为目的节点的概率越大^[66-71]。

上述模型是基于节点拥塞的流量模型,在这种流量模型下网络优化的瓶颈是节点的有限的处理能力。一般假设所有节点的处理能力都相同,即 $C=1$ 。然而,在许多实际网络中,比如因特网,网络优化的瓶颈也存在于链路上,即受有限的链路带宽的限制。因此,本书将主要采用基于节点拥塞的流量模型,也采用基于链路拥塞的流量模型。在基于链路拥塞的流量模型中,一般认为网络中的链路是有向的,将无向网络中的每条无向链路看作是两条反向的有向链路,每条链路的带宽为一个有限值 D 。每条有向链路的输出接口处有一个不受队列长度限制的缓存队列,数据包在队列中采用先进先出的排队规则,每条链路的带宽有限,基于链路拥塞的有序参数可描述为:

$$H(R) = \lim_{t \rightarrow \infty} \frac{D \langle \Delta W \rangle}{R \Delta t} \quad (1-4)$$

在基于链路拥塞的流量模型中,每个节点的处理能力一般等于与其相邻的所有链路的处理能力之和。

1.2.3 主要参数

1.2.3.1 度及度分布

网络中节点的度 (degree) 是节点非常重要的属性,表示该节点具有的直接邻居节点的个数。一般情况下,节点 i 的度用 k_i 或 $k(i)$ 来表示。网络中节点的度分布 (degree distribution) 一般用函数 $P(k)$ 来表示,它表示对于一个随机指定的节点,其度值等于 k 的概率,直观理解为网络中度为 k 的节点数与节点总数的比值。因此,网络中度分布函数在统计意义上表示为:

$$P(k) = \frac{n_k}{N} \quad (1-5)$$

式中： n_k 表示网络中度为 k 的节点个数， N 表示网络中节点总数（或被称为网络规模）。

1.2.3.2 介数

介数（betweenness centrality）表示节点在网络中的中心性，可以用来从理论上估计网络中所有节点的相对负载情况。通俗地讲，某个节点的介数表示通过该节点的最短路径的条数，公式描述为^[72-76]：

$$B_v = \sum_{s \neq d} \frac{\sigma_{sd}(v)}{\sigma_{sd}} \quad (1-6)$$

式中： B_v 表示节点 v 的介数， σ_{sd} 表示从源节点 s 到目的节点 d 的最短路径的条数， $\sigma_{sd}(v)$ 表示从源节点 s 到目的节点 d 的最短路径中经过节点 v 的最短路径的条数。

介数的定义基于最短路径路由算法，为了在非最短路径路由算法下也能够用介数来评估在某种实际路由算法下节点的负载情况，很多学者将介数作了扩展，提出有效介数的概念，有效介数一般定义为：

$$B_v^{\text{eff}} = \sum_{s \neq d} \frac{\sigma'_{sd}(v)}{\sigma'_{sd}} \quad (1-7)$$

式中： B_v^{eff} 表示节点 v 的有效介数， σ'_{sd} 指在某种路由算法下从源节点 s 到目的节点 d 的路径条数， $\sigma'_{sd}(v)$ 表示从源节点 s 到目的节点 d 的路径中通过节点 v 的路径条数。

可以用介数对网络传输容量作一个理论评估。每个时刻，到达网络中任何一个节点 i 的平均数据包个数可以表示为 $RB_i / N(N-1)$ ，其中 $B_i / N(N-1)$ 表示一个数据包到达介数为 B_i 的节点 i 的概率。如果 $RB_i / N(N-1) > C$ ，那么数据包将在该节点 i 上堆积，网络拥塞将会发生。因此，为了保证网络中所有的节点都不

发生数据包堆积，必须保证到达网络中节点介数最大的节点上的数据包数目不超过其处理能力，即必须满足 $RB_{\max} / N(N-1) \leq C$ ，即：

$$R \leq \frac{CN(N-1)}{B_{\max}} \quad (1-8)$$

式中： $B_{\max}$ 表示网络中所有节点介数中的最大值。网络传输容量 R_c 为使得网络不发生拥塞的最大 R 值，则

$$R_c = \frac{CN(N-1)}{B_{\max}} \quad (1-9)$$

从式1-9可以看出，网络传输容量 R_c 与网络中节点介数的最大值 $B_{\max}$ 成反比。因此，在研究中可以用网络中节点介数的最大值 $B_{\max}$ 的变化来评估在不同优化策略下网络传输容量 R_c 的变化情况。

类似于节点的介数，网络中边的介数也可以用来评估边（或链路）上的相对流量负载，边 l_{ij} 的介数表示经过该边的最短路径的条数，公式描述为：

$$B_{ij} = \sum_{s \neq d} \frac{\sigma_{sd}(l_{ij})}{\sigma_{sd}} \quad (1-10)$$

式中： B_{ij} 表示边 l_{ij} 的介数， σ_{sd} 表示从源节点 s 到目的节点 d 的最短路径的条数， $\sigma_{sd}(l_{ij})$ 表示从源节点 s 到目的节点 d 的最短路径中经过边 l_{ij} 的最短路径的条数。

也可以将边的介数进行扩展，边的有效介数（或有效边介数）可以被定义为：

$$B_{ij}^{\text{eff}} = \sum_{s \neq d} \frac{\sigma'_{sd}(l_{ij})}{\sigma'_{sd}} \quad (1-11)$$

式中： B_{ij}^{eff} 表示边 l_{ij} 的有效介数， σ'_{sd} 指在某种路由算法下从源节点 s 到目的节点 d 的路径条数， $\sigma'_{sd}(l_{ij})$ 表示从源节点 s 到目的节点 d 的路径中通过边 l_{ij} 的路径的条数。

在基于链路拥塞的流量模型中，网络优化的瓶颈是链路。类似于基于节点拥