

Internet 高性能组播路由算法研究^{*})

周 灵^{1,2} 孙亚民¹ 卢先领¹

(南京理工大学计算机科学与技术系 南京 210094)¹ (长沙大学电子与通信工程系 长沙 410003)²

摘 要 为了进一步适应 Internet 网络多媒体业务应用, 高性能 Internet 组播路由算法已成为国内外网络研究热点和难点问题之一。分析了组播路由问题的定义和分类; 论述了现有的 IP 层组播路由算法, 主要包括五种类型: 最短路径树算法、最小生成树算法、Steiner 树算法、单约束的 Steiner 树算法和多约束的 Steiner 树算法; 并对它们进行比较和评价。最后提出了高性能 Internet 组播路由算法具有的特点和进一步的研究方向。

关键词 Internet, 组播路由, 高性能, 综述

A Survey of Internet High-Performance Multicast Routing Algorithm

ZHOU Ling^{1,2} SUN Ya-Min¹ LU Xian-Ling¹

(Dept. of Computer Sci. & Tech., Nanjing University of Sci. & Tech., Nanjing 210094)¹

(Dept. of Electronic & Communication Engineering, Changsha University, Changsha 410003)²

Abstract In order to accelerate the development of multimedia application in Internet, high-performance multicast routing algorithm has been discussed and is becoming an important and difficult research problem in Internet. First multicast routing is defined and categorized in the paper. Secondly, all important multicast routing algorithms are depicted and discussed according to different tree kinds: the SPT algorithm, the MST algorithm, the Steiner tree algorithm, the constrained Steiner tree algorithm and the multi-constrained Steiner tree algorithm. At the same time the paper evaluates and compares them with performance and computing complex. At last the character of high-performance multicast routing algorithm and the future research work is proposed.

Keywords Internet, Multicast routing, High-performance, Survey

1 引言

随着计算机、网络技术的飞速发展, 用户对 Internet 网络通信性能提出了更高要求, 对网络带宽的需求呈现直线上升趋势。特别近来各种多媒体网络应用业务的进一步扩展, 视频点播、网络会议、IP 电话、远程教育、交互游戏、分布式多媒体数据库等大量新兴多媒体通信业务急速增加, 导致网络带宽的消耗和网络拥挤日见显著, 下一代 Internet 高性能的研究和应用已经开展如火如荼。这类业务有共同的特点: 一是数据量大, 需要消耗大量的网络带宽; 二是在网络应用上有较严格的服务质量 (Quality of Service, 简称 QoS) 约束条件; 而且, 通常采用一对多 (或者多对多) 的数据通信方式。传统的点到点单播通信方式或广播方式已无法满足当前网络信息传输的要求, 不仅浪费大量的网络带宽, 而且效率很低。显然从路由方面考虑 IP 组播 (Multicast) 路由是最佳的解决方式^[1]。

在组播通信中, 发送方只要发送一个信息包, 所有目的地将收到同一信息包, 每个链路上只有一个信息包的拷贝, 这样大大节省了网络带宽, 同时实现了一对多 (或者多对多) 的数据通信方式。因此组播是最适应 Internet 高性能多媒体网络应用的一种通信方式。通过采用恰当的组播通信策略是可以满足各种多媒体业务约束条件的, 主要是看采用何种方式实现组播。实践证明, 实现组播通信的关键是组播路由算法的实现^[2], 即如何构建一棵简单、有效、健壮的组播路由树。

2 组播路由问题

2.1 组播路由问题的定义

计算机网络通信包括单播、组播和广播三种网络通信类型。组播相对单播而言的, 指一个节点同时向多个目的节点发送信息的通信方式。

在 Internet 中单播路由指寻找从信息源到一个目的节点的满足条件的信息转发路径; 组播路由相对单播路由而言, 就是寻找从信息源到一组 (多个) 目的节点的满足条件的信息转发路径, 通常以构造一棵以源为根的信息转发树的形式出现。组播路由可以看作是路由的一种普遍的形式, 若目的节点为一个节点, 则组播路由即变为单播路由; 若目的节点为源所在网络的所有节点, 则组播路由问题就是广播路由问题。

网络可用二元组 $G(V, E)$ 来表示, 其中 V 为网络节点集, E 为网络的边集合, 则网络节点数为 $n = |V|$, 给应组播信息源 s , 目的节点集点 M , 目的节点数为 $m = |M|$, 对 $e \in E$, 可以定义代价函数 $cost(e): E \rightarrow R^+$; 网络状态函数例如时延函数 $delay(e): E \rightarrow R^+$ 。给定一个组播组 M 和一组可能的优化目标函数 O , 组播路由就是以网络拓扑结构和网络状态为基础, 构建一棵使目标函数达到最优的组播树 T 的过程。其主要元素与关系可以用图 1 表示如下^[3]; 约束集 C 通常以时延 Delay、端端间时延抖动 Delay Jitter、最小带宽 Bandwidth、包丢失率 Packet Loss Rate (PLR) 或者它们的组合形式出现。

^{*}) 本文受到湖南省教育厅科研基金项目资助 (项目编号: 05C059)。周 灵 博士生、讲师, 主要研究领域为计算机网络通信, 路由算法与网络协议; 孙亚民 教授、博士生导师, 主要研究领域为计算机网络通信, 网络算法与协议; 卢光领 博士生、讲师。

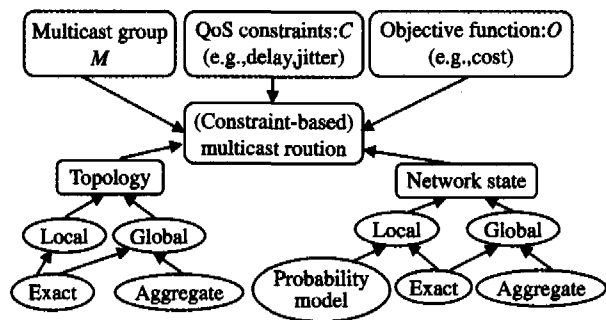


图1 组播路由的组成

2.2 组播路由问题分类

目标函数优化包括树优化和链路优化两种类型^[4]。树优化通常以最小化组播树的代价(Cost)的形式定义,代价可以是整个组播树使用的带宽或者其他网络利用率的单调非减函数。链路优化就是使用链路优化函数构造组播树的最优路由链路,例如,使组播树中使用的链路带宽最大。

约束包括两种类型:链路约束和树约束^[4]。链路约束是对路由选择所使用的链路的限制。例如,链路约束要求链路上的带宽或可用缓冲大于或等于一个预定义的值。树约束包括:(i)对组播树中从源端到接收端的每条路径上性能量度的组合值的限制。例如,从源端到所有接收端路径上的端到端时延限制。(ii)对从同一源端到任意两个接收端路径上的性能量度的不同组合值的限制。例如,接收端之间的时延抖动限制定义为从相同源端到任意两个不同接收端路径上的端到端时延差异。在不同的QoS情况下,每个接收端可能会有不同的约束。下面表1描述了各种约束和优化的组合情况,(1)~(12)表示各类组播路由问题。

表1 组播路由问题的分类

	无优化	链路优化	树优化
无约束		(6)链路优化问题	(7)树优化问题
链路约束	(1)链路约束问题	(8)链路约束	(9)链路约束树
	(2)多链路约束问题	链路优化问题	优化问题
树约束	(3)树约束问题	(10)树约束链	(11)树约束树
	(4)多树约束问题	路优化问题	优化问题
链路和树约束	(5)链路及树约束问题		(12)链路及树约束树优化问题

应用中比较常见的组播路由问题主要有7种类型:

链路约束问题:规定链路约束来构造可行的组播树。例如,带宽约束的组播路由。

树约束问题:规定树约束来构造可行的组播树。例如,时延约束的组播路由。

链路优化问题:使用一个链路优化函数来构造一个优化的组播树。例如,组播树中链路上的带宽最大的组播路由。

树优化问题:使用树优化函数来构造一个优化的组播树。例如,组播树的总代价最小。这也就是著名的Steiner树问题。

链路约束树优化问题:规定链路约束并使用一个树优化函数来构造一个优化的组播树并实现约束。例如,带宽约束Steiner树问题。

树约束树优化问题:使用树约束和树优化函数来构造一个优化的组播树。例如,时延约束Steiner树问题。

链路约束及树约束树优化问题:使用链路和树约束和树

优化函数来构造一个优化的组播树。例如,带宽和时延约束的树优化问题。

3 组播路由算法

在网络通信中主要是通过组播路由算法构造组播树来解决组播路由问题,除极少数情况下,大多数组播路由树的构造都是NP-hard问题;因此一般通过设计启发式组播路由算法来寻找一棵次优组播树。下面对各种组播路由算法从所使用的目标函数和约束条件方面作分类综述,主要包括五种类型:最多路径树算法;最小生成树算法;Steiner树算法;单约束的Steiner树算法和多约束的Steiner树算法。

3.1 最短路径树算法(Short Path Tree,简称SPT)

最短路径树是从组播源到每个目的节点的路径都是两者之间的最短路径的组播树。如果使用跳数作为量度,则最短路径树就是最小跳数树。如果用时延作为量度,则最短路径树就是最小时延树。最短路径树算法可以解决树约束问题(例如时延约束问题),而且都是在多项式时间内能够求出问题的精确解。

目前Internet中广泛使用的SPT算法是著名的Dijkstra算法^[5]和Bellman-Ford^[6]算法。Dijkstra算法主要运用路径递增的思想贪心地构造一棵最短路径树,是一个基于源的集中式算法;Bellman-Ford算法是一个不错的分布式最短路径算法;但两者均不考虑组播树的代价优化。近来的DDSP算法^[7]是Dijkstra SPT的扩展算法,构造最短路径时借鉴目的节点驱动的思想通过目的节点共享路径,从而达到降低SPT树的总体代价的目的。SBPT算法^[8]则通过目的节点的动态加入时,不断地在邻居节点中寻找最短路径,从而构造一棵低代价的SPT树,是一个动态的SPT树算法。

3.2 最小生成树算法(Minimum Spanning Tree,简称MST)

最小生成树是连接网络中所有节点并且树的代价最小的一棵组播树。典型算法是Prim算法和Kruskal算法^[9]。两者都能在多项式时间内能够求出问题的精确解,可以解决树优化组播问题。

3.3 Steiner算法

Steiner树问题是组播路由中的一个经典而重要的问题。若所有节点都是组播目的节点时Steiner树问题就是最小生成树问题,但大多情况下Steiner问题是要求连接网络中部分节点的最小代价树,这是一个NP-hard问题,没有多项式时间的算法可以解决树优化组播问题。经典的算法有KMB算法^[10]、MPH^[11]算法、ADH算法^[12]等。

KMB算法是由Kou等人提出的求解Steiner组播树算法,其复杂度为 $O(mn^2)$;ADH(Average Distance Heuristic)算法复杂度为 $O(n^3)$;MPH算法复杂度证明也为 $O(mn^2)$,但通常情况下生成的树费用要小于KMB生成树的费用;SPH是Doar提出的以源为根的最小费用路径算法,是最简单的Steiner启发式算法但树代价优化性能较差。到目前为止,MPH仍是一个非常优秀的求解Steiner树问题的启发式算法。

3.4 单约束的Steiner树算法

单约束的Steiner问题是将Steiner问题加上单约束如时延约束、时延抖动约束等。研究最多的是时延约束的最小代价树(时延约束的Steiner树)。可以解决单约束的树优化组播问题。目前主要算法有BSMA算法^[13]、KPP算法^[14]、

CDKS算法^[15]等。另外 Jia 也提出了一个时延约束的分布式算法^[16]。

BSMA 算法由 Zhu 等人提出的一个时延受限的最小代价组播启发式算法,主要利用 Dijkstra 生成最小时延树,再利用“超边”迭代减少组播树的代价(其代价更小的超边是使用 k 最短路径算法计算的),其计算复杂性为 $O(kn^3 \log n)$,其中 k 为超边迭代次数。KPP 算法由 Kompella 等人提出,实质是扩展 KMP 到有时延约束的情况,首先计算图 G 满足时延约束的闭包图 G' ,然后用 Prim 计算 G' 的最小生成树在还原到 G 中的边,计算复杂度为 $O(\Delta n^3)$,其中 Δ 为时延约束。CDKS 算法是 Sun 等提出的一个启发式算法,主要是分别用费用和时延为度量计算最短路径树,然后合并两棵树来满足时延的同时降低代价,复杂度较低为 $O(n^2)$,但生成树的代价较高。

DCMPH(the Delay-Constrained Minimum Path Heuristic)是我们根据 MPH 提出了一个时延受限的最小代价组播启发式算法,有很好的代价值能和较低的计算复杂度。

3.5 多约束的 Steiner 树算法

多约束的 Steiner 树算法指生成树受到多个条件的约束(例如同时受时延和时延抖动约束等约束)条件的最小代价组播树算法,可以解决多约束的树优化组播问题。目前算法有 Rouska 等提出的同时受时延和抖动约束的 DVMA 算法^[17]、王珩提出的模拟退火解决时延和时延抖动约束的 SADVNA 算法^[18]、孔令山等提出的多约束组播路由算法^[19]等。

下面将比较重要的组播路由算法各方面性能进行比较总结如下,详见表 2。

表 2 组播路由算法比较

算法类别和名称		集中式/分布式	发起方	复杂度	约束/优化问题
最短路树算法	Dijkstra	集中式	源节点	$O(n^2)$	树约束
	DDSP	集中式	源节点	$O(n^2)$	树约束
	LCSPT	集中式	源节点	$O(n^2)$	树约束
最小生成树	Prim	集中式	源节点	$O(n^2)$	树优化
	Gallager	分布式	接收方	$O(n \log 2n)$ 消息复杂度 $O(n \log 2n + m)$	树优化
Steiner 树算法	KMB	集中式	源节点	$O(mn^2)$	树优化
	MPH	集中式	源节点	$O(mn^2)$	树优化
单约束 Steiner 树算法	BSMA	集中式	源节点	$O(kn^3 \log n)$	时延约束树优化
	KPP	集中式	源节点	$O(\Delta n^3)$	时延约束树优化
	Jia	分布式	源节点 接收方	$O(2m)$ 消息复杂度 $O(2m)$	时延约束树优化
多约束 Steiner 树算法	Rouskas	集中式	源节点	$O(klmn^4)$	时延、时延抖动约束
	HMCMC	集中式	源节点	$O(m(n \log n + m))$	多 QoS 约束

* 算法复杂度 n 为网络节点数, m 为目的节点数, e 为网络的链路数。其他 k, l, Δ 等为具体算法的特定参数。

根据不同的分类标准组播路由算法有不同的分类方法。根据是否考虑目的节点的动态加入/退出可将组播路由算法分为静态组播路由算法和动态组播路由算法,静态算法当目的节点发生变化或者网络拓扑结构/状态变化时算法不能动态适应,需要重新计算路由树。动态路由算法能根据节点的动态变化来动态计算新的组播树,不干扰已经存在的组播树的通信。因为不需要重新计算整个组播树,因而计算复杂度相对也较低。动态组播路由根据每次成员节点加入/离开后组播树是否需要重组又可分为三类:不重组的路由算法、部分重组的路由算法和完全重组的路由算法^[20]。如果对动态组播进一步考虑 QoS 的约束的话,则成为受 QoS 约束的动态组播路由问题。目前组播算法多为静态算法。

根据网络拓扑信息和链路状态信息的收集和组播树的计算方式还可分为集中式的路由算法和分布式的路由算法。集中式算法每个节点需要保持整个网络拓扑和状态,由中心节点(通常是源节点)集中计算组播树,然后根据生成的树转发组播信息。其缺点式中心节点负荷大,容错性不强;计算量比较大,不适合实时路由算法;不利于高性能网络的可扩展性。分布式算法则每个参与计算的节点只需要知道局部的网络拓扑和状态信息,各自独立地计算出加入组播树的路径进一步合并为一棵组播树。显然分布式组播路由克服了集中式路由的缺点。Jia 等已经提出了一些分布式组播算法,但远远不如

集中式算法研究的深入^[21]。

4 Internet 高性能组播路由算法研究热点和发展趋势

首先,Internet 高性能组播路由算法对 Internet 动态性的支持。对于网络拓扑变化和网络状态信息变化应该由动态算法来动态处理,以及目的节点的动态变化不能仅仅依靠频繁的组播树的重组计算来适应。我们需要根据 Internet 网络组播通信量和组播成员的动态的特性设计适合 Internet 高性能组播路由算法。目前静态组播路由算法由于没考虑网络的动态性,很多情况下仅具有理论上的意义。

其次,Internet 高性能组播路由算法对分布式计算方式的支持。集中式算法计算量大,很难在实际网络中应用,目前主要是应用 Dijkstra SPT 和 Bellman-Ford SPT 算法,其复杂度在组播算法中是最低的,一般情况下仅仅为 $O(n^2)$,但相对当前的网络状态计算复杂性已经很高了;而且考虑多 QoS 约束的组播算法复杂度应通常在 $O(n^3)$ 以上,难以应用到实际的 Internet 路由中。分布式算法能相对分散中心点计算任务,降低了算法的计算复杂度。

第三,Internet 高性能组播路由算法在计算路由时如何实现 Internet 网络拓扑的分层结构问题,和对网络链路状态信息如何实现信息聚合而尽量少丢失信息的问题。

第四,提高生成组播树的代价性能和降低算法计算复杂度的折中也是高性能组播路由算法需要研究的核心问题之一。算法的代价性能和计算复杂度通常是一对矛盾的指标。要降低算法复杂度必然影响算法性能的提高,而要有好的代价性能则算法复杂度就会增加。一个好的组播算法总是要在组播树的总体代价与计算复杂度之间做出权衡,这仍然是 Internet 高性能组播路由算法要研究的一个问题。

最后,适应下一代 Internet 高性能网络的各种特定 QoS 要求的组播算法也是研究的一个方面。下一代 Internet 组播业务众多,不同的业务有不同的需要,而且服务质量要求更加严格,满足各种特定要求的 QoS 约束组播路由算法仍是研究的一个重要方面。

小结 通过对 Internet 组播路由问题的描述和国内外组播路由算法研究现状的分析,可以得出高性能组播路由算法是网络通信研究的一个难点和热点问题。组播路由的目标是寻找一覆盖源节点和所有目的节点的组播树,而这棵组播树必须由一个高效的组播路由算法来生成实现;因此组播路由算法就是要在通信网中建立一棵以信息源为根简单、有效、健壮的信息转发树。目前 QoS 约束的静态组播路由算法已经有了比较深入的研究,下一阶段着重研究 Internet 高性能组播路由算法,主要表现为动态、分布式、分层/聚合、高代价性能、低复杂度、多 QoS 约束的组播路由算法。以后将以以上讨论为背景,对 Internet 高性能组播路由算法展开研究。

参考文献

- Holbrook H, Cain B. Source-Specific Multicast for IP. draft-ietf-ssm-arch-06. Mar 07, 2005
- Laxman H, Sahasrabudhe, Biswanath Mukherjee. Multicast Routing Algorithms and Protocols; A Tutorial, IEEE Network, 2000(1):92~102
- Wang Bin, Hou J C. Multicast Routing and its Extension; Problems, Algorithms, Protocols [J], IEEE Network, January 2000 (1):22~36
- 徐格,吴建平,徐明伟. 高等计算机网络. 北京:机械工业出版社, 2003

(上接第 15 页)

了体全息存储记录通道噪声模型。依据此模型,可以设计适当的信源编码和信道编码技术来抑制存储系统中噪声的影响,为体全息存储技术实用化奠定基础。

参考文献

- Burr G W. Holographic storage. Encyclopedia of Optical Engineering. In: Johnson R B, Driggers R G, ed. New York: Marcel Dekker, 2003. 1~15
- Hong J F, McMichael I, Chang T Y, et al. volume holographic memory systems; techniques and architectures. Opt Eng, 1995, 34(8):2193~2203
- Coufal H J, Psaltis D, Sincerbos G T. Holographic data storage. In: Berlin, Germany: Springer-Verlag, 2000. 3~87, 283~479
- Psaltis D, Burr G W. Holographic data storage. IEEE Computer, 1998, 31(2): 52~60
- 陶世荃主编. 光全息存储. 北京:北京工业大学出版社, 1998
- Wu Fei, Xie Chengsheng, Liu Zhaobin. Signal for the volume holographic storage system. Computer and Communication, 2005, 4
- Gerritsen H J, Hannan W J, Ramberg E G. Elimination of speckle noise in holograms with redundancy. Appl Opt 1968, 7(11): 2301~ 2311
- Amako J, Miura H, Sonehara T. Speckle-noise reduction on kinoform reconstruction using a phase-only spatial light modulator. Appl Opt, 1995, 34(17): 3165~3171
- Rajenbach H, Pelboulbe A, Huignard J P. Noise suppression in

- Dijkstra E. A Note on Two Problems in Connection with Graphs [J]. Numerical Mathematic, 1959(1): 269~271
- Andrew S. Tanenbaum. Computer Networks (Third Edition). Prentice Hall International, Inc. 1996
- Zhang Baoxian, Mouftah H T. A Destination-Driven Shortest Path Tree Algorithm. In: Proc. of the 2002 IEEE Int'l Conf. on communication, 2002(4): 2258~2261
- Fujinoke H, Christensen K J. The New Shortest Best Path Tree (SBPT) Algorithm for Dynamic Multicast Trees. In: Proc. of the 24th IEEE Int'l Conf. on Local Computer Networks, 1999. 204~210
- Cormen T H, Leiserson C E, Rivest R L. Introduction to Algorithms. MIT Press, 1997
- Kou L, Markowsky G, Berman L. A fast algorithm for Steiner trees in graphs. Acta Informatica, 1981, 15(2): 141~145
- Winter P. Steiner problem in networks; survey [J]. Networks, 1987, 3:129~167
- Rayward-Smith V J. The computation of nearly minimal steiner trees in graphs [J]. Int J Math Educ Sci Technol, 1983, 14:15~23
- Zhu Q, Parsa M, Garcia-Luna-Aceves J J. A source-based algorithm for delay-constrained minimum-cost multicasting. In: Proc. IEEE INFOCOM'95, 1995. 377~385
- Kompella V P, Pasquale J C, Polyzos G C. Multicasting for multimedia applications. In: Proc. IEEE INFOCOM'92, 1992. 2078~2085
- Sun Q, Langendoerfer H. Efficient multicast routing for delaysensitive applications. In: Proc. Second Workshop Protocols Multimedia Systems (PROMS'95), Oct. 1995. 452~458
- Jia Xiaohua. A Distributed Algorithm of Delay-Bounded Multicast Routing for Multimedia Applications in Wide Area Networks. IEEE Transaction on Networking, 1998, 6(6): 828~837
- Rouskas R N, Baldine I. Multicast routing with end-to-end delay and delay variation constraints. IEEE Journal on Selected Areas in Communications, 1997, 15(3): 346~356
- 王珂, 孙亚民. 一种基于拉格朗日松弛的时延约束组播路由算法 [J]. 通信学报, 2004, 25(5): 92~102
- 孔令山, 丁炜. 一种多约束 QoS 多播路由算法. 通信学报, 2003, 24(7): 30~36
- Raghavan S, Manimaran G, et al. A Rearranging algorithm for the construction of delay-constrained dynamic multicast trees. IEEE Transaction on Networking, 1999, 7(4): 514~529
- Salama H F. Evaluation of Multicast Routing Algorithm for Real-Time Communication on High-Speed Networks [J]. IEEE Journal on Selected Areas in Communication, 1997, 15(3): 332~345

photorefractive image amplifier. Opt Lett, 1989, 14(22): 1275~1277

- Rabinovich W S, Feldman B J, Gilbreach G C. Suppression of photorefractive beam fanning using achromatic gratings. Opt Lett, 1991, 16(15): 1147~1149
- Asthana P, Nordin G P, Tanguay A R, et al. Analysis of weighted fan-out/fan-in volume holographic optical interconnection. Appl Opt Lett, 1993, 32(8): 1441~1469
- Gu C, McMichael I, Saxena R, et al. Cross-talk-limited storage capacity of volume holographic memory. Opt Soc Am A, 1992, 9(11): 1978~1983
- Curtis K, Gu C, Psaltis D. Cross talk in wavelength-multiplexed holographic memories. Appl Opt Lett, 1993, 18(12): 1001~1003
- Keskinoz M, Kumar B V K V. Application of linear minimum mean-squared-error equalization for volume holographic data storage. Appl Opt, 1999, 38(20): 4387~4393
- Vadde V, Kumar B V K V. Channel Estimation and Intra-page Equalization for Digital Volume Holographic Data Storage. SPIE, 3109: 245~250
- Vadde V, Kumar B V K V. Efficient Modeling of Volume Holographic Storage Channels (VHSC). IEEE, 2000. 116~118
- Vadde V, Kumar B V K V. Efficient Modeling and Iterative Magnitude-Squared Decision Feedback Equalization (DFE) for volume holographic storage channels. IEEE, 2001. 2696~270
- 陈端容. 全息存储通道中 6:8 等调制码的优化研究 [D]. 见: 武汉:华中科技大学. 2003, 11: 22~24
- 刘思敏, 许京军, 郭儒. 相干光学原理及应用. 见: 天津:南开大学出版社, 2001. 52~61