

光分组交换网络的冲突解决： 用光纤延迟线还是波长转换器¹

张治中* ** 曾庆济* 王建新* 维江涛* **

*(上海交通大学区域光纤通信网与新型光通信系统国家重点实验室 上海 200030)

** (重庆邮电学院通信工程学院 重庆 400065)

摘 要: 该文研究了光分组网络的冲突处理问题, 提出了两种异于传统设计方法的交换结构。研究表明: 对于非突发性业务, 非简并排列方式的光纤延迟线是成本最低的解决手段; 而对于突发业务, 将波长转换器和延迟线结合使用是兼顾分组丢弃性能, 系统体积和成本的较好策略。文中基于此给出了系统设计原则, 该原则可确保系统节约大量的波长转换器和延迟线, 对改善系统的整体性能十分有利。

关键词: 光分组交换, 光缓存, 波长转换器, 共享, 分组丢弃率

中图分类号: TN929.1 **文献标识码:** A **文章编号:** 1009-5896(2004)11-1812-07

Contention Resolution in Optical Packet Switched Networks: Using Fiber Delay-Lines or Wavelength Converters

Zhang Zhi-zhong* ** Zeng Qing-ji* Wang Jian-xin* Luo Jiang-tao* **

*(*R&D Center for Broadband Optical Networking Tech.,
Shanghai Jiaotong Univ., Shanghai 200030, China*)

** (*Dept of Commun. Eng., Chongqing Univ. of Posts and Telecomm.,
Chongqing 400065, China*)

Abstract This paper proposes two new switch architectures that are more flexible and cost effective than the methods introduced in the literatures. For a given packet loss probability requirement, the results demonstrate that (1) for non-bursty traffic, non-degenerate form of Fiber Delay-Lines (FDLs) is the most cost effective solution; (2) for bursty traffic, a combined use of degenerate FDLs and Tunable Wavelength Converters (TWCs) is a cost effective and robust solution. With the increase of the average burst length, i.e., burstiness, the number of TWCs needs to be increased so as to maintain a reasonable packet loss probability. However, even for the traffic with high degree of burstiness, the architecture that employs a set of degenerate FDLs and TWCs shared among the input lines, is still a cost effective and robust solution.

Key words Optical packet switching, Optical buffering, Wavelength converter, Sharing, Packet loss probability

1 引言

光分组交换属于分组级的光信号处理, 与光时分复用 (OTDM) 技术相比, 对光器件工作速度的要求大大降低, 与波分复用 (WDM) 技术相比, 能以更细的粒度快速分配信道, 更加灵活、有效地利用带宽。因此, 光分组交换网一经提出便受到广泛关注, 并得到大量研究^[1-7]。

¹ 2003-05-15 收到, 2004-01-07 改回

国家自然科学基金 (69990540) 和国家 863 计划 (2001AA121073) 资助项目

在光分组交换网络中，当两个或更多的分组要求同时离开交换节点的同一输出端口时，会产生冲突。若没有采取相应的措施，将造成分组丢失。基于此，人们提出了各种方案，以解决该问题，可以概括为^[2-5]：(1) 从时间域上解决冲突，最常用的方法是采用光纤延迟线 (Fiber Delay Line, FDL) 缓存分组，直到所争用的波长通道空闲为止；(2) 从波长域上解决冲突，即通过波长变换将冲突分组的承载波长变换到另一空闲波长上，使分组在这一空闲波长通道上传送到下一节点；(3) 从空间域上解决冲突，即偏转路由方式，利用网络中的闲置路由，使分组绕过冲突节点迂回去往目的节点。

偏转路由方式是将网络中的其它链路和节点作为“缓存器”来解决冲突的，它所引起的接收分组乱序会很严重，还会带来较大的延时和时延抖动。当网络负载较重或网络连接度较低时，偏转路由方式不但不能有效地解决冲突，还会恶化全网的吞吐量。因此，人们更关注于从时间域和波长域两方面来解决冲突。

在波长域方面，Danielsen^[6,7] 等人最早研究了交换节点纯粹利用波长转换器解决冲突时的分组丢弃性能，基于非突发业务的仿真结果表明：利用目前较成熟的 WDM 技术，节点可以不用 FDL 来解决冲突。但该方法所用的转换器太多，达 NW 个 (N 为输入 / 输出光纤端口数， W 为每纤的波长数)。为了克服该不足，Eramo^[3,8] 等人提出将波长转换器设计成能为每个输入信道共享的形式，该结构比 Danielsen 等人提出的结构节约了数量可观的转换器，但交换矩阵所用的光门数有一定的增加。

我们近期基于时间域和波长域的研究表明^[9]：(1) 对于非突发业务，同样数量的 FDL 比波长转换器能更有效地解决冲突，且便宜得多。因此，Danielsen 和 Eramo 等人所提的方案在非突发业务环境下，并不是经济有效的；(2) 对于突发业务，特别是中高级突发度的业务，同样数量的波长转换器比 FDL 能更有效地解决冲突，在系统体积上也能得以减少。

综上讨论，我们是否可以将 FDL 和波长转换器结合使用，在满足分组丢弃性能的条件下，在系统性能，成本和体积间取得较为折衷的设计方案呢？这正是本文将要解决的问题。

2 交换结构及分析

2.1 交换节点模型

图 1 示出了本文的交换节点模型 (冲突解决模块在后面介绍)，它具有 N 个基于 WDM 技术的光纤输入 / 输出端口，每根光纤支持 W 个波长，记为 $\{\lambda_1, \dots, \lambda_w\}$ 。如图 1 所示，交换结构由 3 部分构成：(1) 输入级，由一组光波分解复用器组成。解复用器将各个输入光纤上的波长解复用后，送入下一级 (交换级) 进行处理。每个输入 / 输出信道中的分组可由变量 $I(i, \lambda_j) / O(i, \lambda_j)$ 进行描述，其中， $i \{i = 1, \dots, N\}$ 表示输入 / 输出光纤， $\lambda_j \{j = 1, \dots, W\}$ 表示波长；(2) 交换级，由空分交换矩阵组成。在这里，输入分组要么被交换到其对应的目的端口，要么被交换到冲突解决模块；(3) 输出级，由一组复用器组成，它们将分组复用到对应的输出光纤端口，进而去往下一节点。

本文采用了时隙网络，假设：(1) 分组具有固定长度，每个波长信道上的分组到达是基于时隙同步的；(2) 一个时隙代表了单个分组的传输时间，一个时隙的持续时间 T 表示一个时间单元。

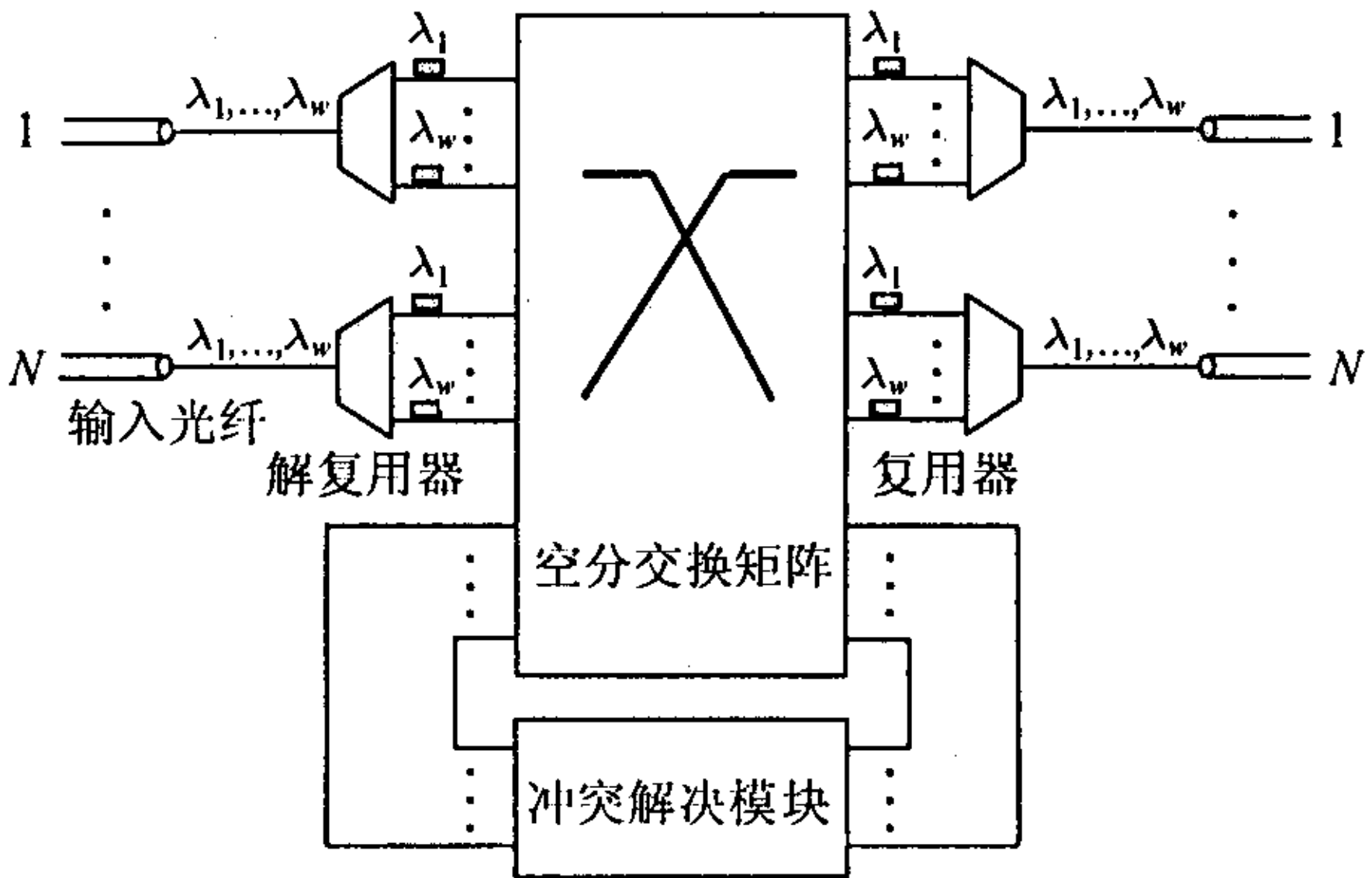


图 1 交换节点模型

2.2 冲突解决模块设计及分析

本文将要对比研究 3 种交换结构，其中第 1 种 (DF 结构) 已为大量文献研究，第 2 和第 3 种为本文提出的结构。它们的不同之处在于冲突解决模块。基于图 1 所示的节点模型，在下文中，我们分别给出了冲突解决模块的内部结构。

反馈共享简并延迟线 (Degenerate FDL, DF) 结构 如图 2(a) 所示，DF 结构的最大特点在于：冲突是通过一组简并形式的延迟线予以解决的 (简并形式：延迟线的缓存时间均匀分布于 T 和 MT 之间)。最短长度的 FDL 对分组提供的缓存时间 (缓存容量) 是 T ，最长的 FDL 提供的缓存时间是 MT ，总的缓存容量为 $(T + 2T + \cdots + MT)$ 。

反馈共享非简并延迟线 (Non-Degenerate FDL, NDF) 结构 NDF 结构如图 2(b) 所示。它也是基于 FDL 解决冲突的，但 FDL 的排列方式与图 2(a) 截然不同，即延迟线的缓存时间并非均匀分布，其中，缓存时间为 iT 的延迟线有 M_i 根 ($i = 1, \cdots, L$)。

由图 2(b) 可见，NDF 结构具有这样的特点：系统总的缓存容量为 $(M_1 + 2M_2 + \cdots + iM_i + \cdots + LM_L)$ 。对比 DF 结构来看，若 $M = (M_1 + M_2 + \cdots + M_L)$ ，即若两种交换结构总的延迟线数量相同时，NDF 所需的延迟线长度要少于 DF 结构。例如，若 $M = 8$ ，且 NDF 只有 3 种长度的延迟线，个数分别为 $(M_1, M_2, M_3) = (3, 3, 2)$ ，则 DF 和 NDF 结构所需的 FDL 长度 (以时隙长度衡量) 分别为 $36T$ 和 $15T$ 。可见，NDF 结构所节约的 FDL 长度是很可观的。因此，系统的体积将大大减少。

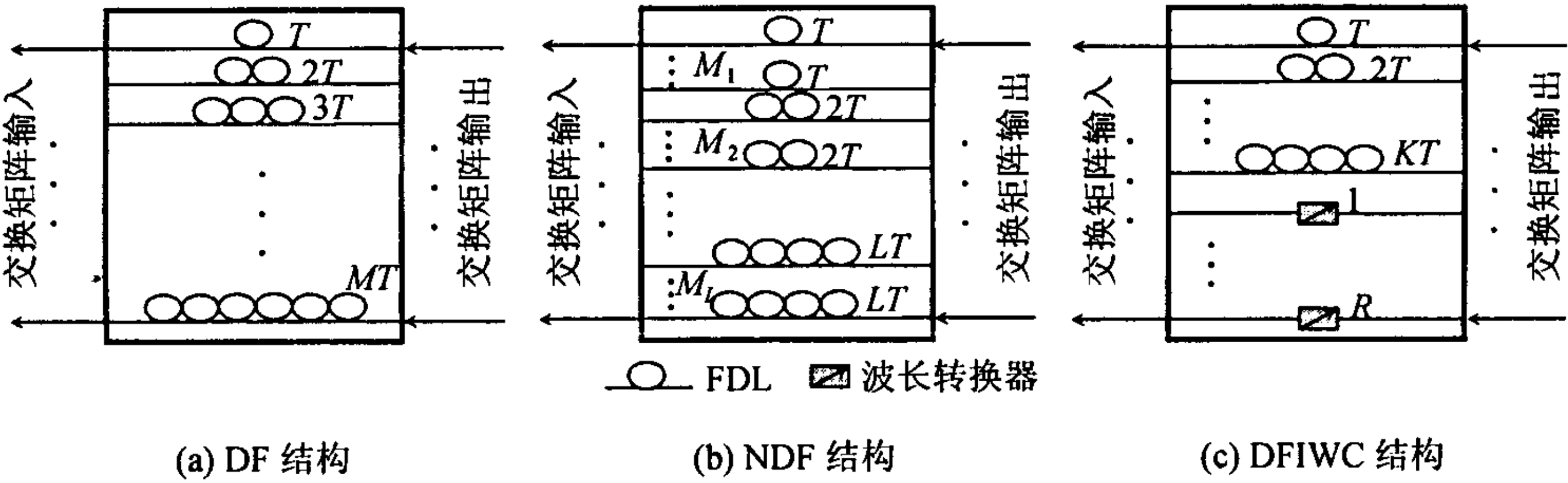


图 2 不同形式的冲突解决模块

反馈共享波长转换器和简并延迟线 (Degenerate FDL and Tunable Wavelength Converters DFTWC) 结构 图 2(c) 给出了 DFTWC 结构的实现方式，该结构中 FDL 的排列是属于简并形式的。除了 K 条 FDL 外，还利用了 R 个波长转换器来解决冲突。对于该结构，若取 $(K, R) = (M, 0)$ ，则 DFTWC 与 DF 结构等效，即 DF 结构是 DFTWC 结构的一个特例。由于该结构更具代表性，我们将对之进行详细的分析。首先定义下列算式：

$$I(i, \lambda_j, t) = \begin{cases} 1, & \text{第 } t \text{ 时隙时, 第 } i \text{ 条输入光纤的波长 } \lambda_j \text{ 有分组到达,} \\ 0, & \text{其他,} \end{cases} \quad i = 1, \cdots, N, j = 1, \cdots, W \quad (1)$$

$$O(i, \lambda_j, t) = \begin{cases} 1, & \text{第 } t \text{ 时隙时, 第 } i \text{ 条输出光纤的波长 } \lambda_j \text{ 有分组离开,} \\ 0, & \text{其他,} \end{cases} \quad i = 1, \cdots, N, j = 1, \cdots, W \quad (2)$$

$$WC(k, t) = \begin{cases} 1, & \text{波长转换器 } k \text{ 在第 } t \text{ 时隙时被占用,} \\ 0, & \text{其他,} \end{cases} \quad k = 1 \cdots, R \quad (3)$$

$$FDL(k, t) = \begin{cases} 1, & \text{光纤延迟线 } k \text{ 在第 } t \text{ 时隙时不可用,} \\ 0, & \text{其他,} \end{cases} \quad k = 1 \cdots, K \quad (4)$$

设第 t 时隙时, 存在有 $I(1, \lambda_1, t) = 1$, $I(N, \lambda_1, t) = 1 (N \neq 1)$, 且两者均指向 $O(1, \lambda_1)$. 若交换调度控制器允许 $I(1, \lambda_1)$ 直接进入 $O(1, \lambda_1)$, 则对于 $I(N, \lambda_1)$, 以下任意一种情况满足时都可以解决冲突:

- (1) 若 $\exists \lambda_j \{j = 2, \dots, W\}$ 及 $k \in \{1, \dots, R\}$, 满足 $O(1, \lambda_j, t) = 0$ 及 $WC(k, t) = 0$;
- (2) 若 $\exists k \in \{1, \dots, K\}$, 满足 $FDL(k, t) = 0$.

第 1 种情况表明, 可以利用波长转换器解决冲突. 第 2 种情况表明, 可以利用 FDL, 将冲突的分组调度到将来的时隙予以传送. 显然, 第 1 种措施对分组传输的时延改善有利, 且使交换矩阵的容量得到更充分利用.

下面, 我们来分析 DFTWC 结构在一种特殊配置下的分组丢弃性能, 即 $(K, R) = (0, R)$. 该情况为下文针对非突发业务的仿真提供了评估上限.

假设每个波长上的分组到达是独立同分布的, 到达分组指向各输出光纤端口的概率均为 $1/N$. 此时系统的分组丢弃率即等价于单个输出端口的分组丢弃率. 令 $p_t = \Pr\{I(i, \lambda_j, t) = 1\}$ 为 t 时隙时输入光纤 i 上的波长 λ_j 有分组到达的概率, 则对于每纤负载为 p 的分组到达, 每信道的负载可表示为

$$p_w = \lim_{t \rightarrow \infty} p_t = \lim_{t \rightarrow \infty} \Pr\{I(i, \lambda_j, t) = 1\} = p/W \quad (5)$$

所有输入光纤中到达的分组同时要求交换到同一输出光纤端口的个数 (h) 服从二项分布:

$$p_a(h) = \left[\begin{matrix} NW \\ h \end{matrix} \right] (p_w/N)^h (1 - p_w/N)^{NW-h} \quad (6)$$

分组的丢弃可能由两种情况引起, 一种是由于转换器的数量不够, 另一种是由于超过了 W 个分组要求同时从同一输出光纤端口输出. 假设转换器数量足够, 则分组丢弃率为

$$p_{\text{loss}} = p_a(h > W) = \sum_{h=W+1}^{NW} \left[\begin{matrix} NW \\ h \end{matrix} \right] (p_t/N)^h (1 - p_t/N)^{NW-h} \quad (7)$$

3 仿真及数值分析

仿真程序由 C++ 实现, 仿真中的每一个数据点都是在 10^9 个分组到达后获得的. 仿真中假设: (1) 输入 / 输出光纤端口数 $N = 16$; (2) 负载为 0.8/ 每纤; (3) NDF 结构最多采用 3 种长度的延迟线, 个数分别为 M_1, M_2 和 M_3 个; (4) DF 和 NDF 结构中的总延迟线数量均为 8 根, 即 $M = (M_1 + M_2 + \dots + M_L) = 8$; (4) DFTWC 结构中的延迟线和转换器总数量之和也为 8, 即 $R + K = 8$; (5) 为减少功率损失, 每个分组最多被循环调度 2 次.

对于非突发业务, 我们假定: (1) 数据分组以贝努利 (Bernoulli) 过程独立同分布到达; (2) 到达分组的目的端口均匀分布于各个输出光纤端口.

对于突发业务, 我们采用了 Hass 等人介绍的方法^[1], 其模型如图 3 所示. 该模型包括 3 个无记忆的状态, 状态间的转移概率如图 3 所示. 对于一个波长信道来说, 若在当前时隙无分组到达, 则它处于 “Idle” (空闲) 状态, 在下一个时隙到来时, 该信道会以 p_a 的概率停留在 Idle 状态, 而以 $(1 - p_a)$ 的概率产生分组, 进入 “Burst I” 状态. 对于当前状态为 “Burst I” 的信道, 在下一个时隙到来时, 该信道会以 p_b 的概率停留在 Burst I 状态 (即有分组到达, 目的端口不变), 而以 $(1 - p_b)$ 的概率进入 Idle 或 “Burst II” 状态 (转移概率分别为 $(1 - p_b)p_a$ 和 $(1 - p_b)(1 - p_a)$), 如果进入了 Burst II 状态, 即意味着该信道仍然有分组到达, 但目的端口已经改变了. 信道处于 Burst II 状态下的转移情况与 Burst I 类似, 不再赘述.

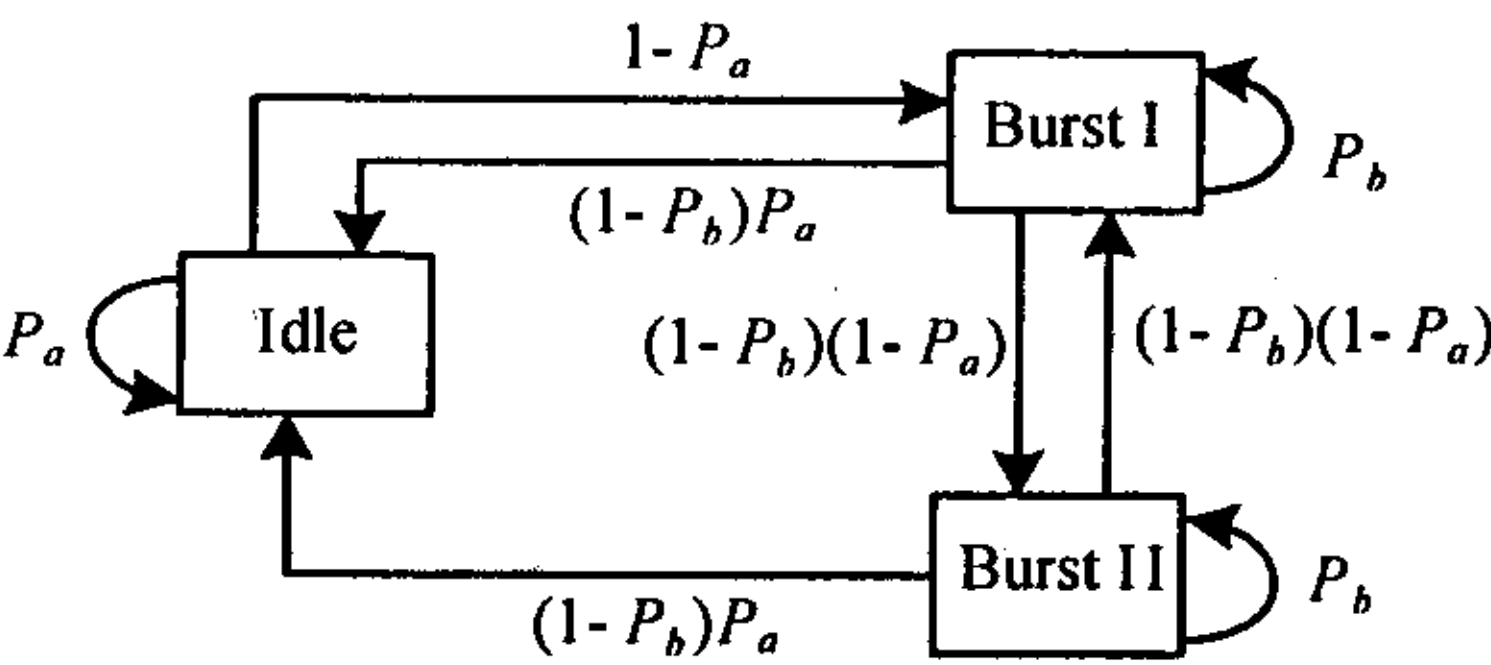


图 3 突发业务产生模型

由以上描述可知，系统产生的突发序列服从二项分布，建立 Markov 平稳方程可求出每信道到达分组的平均长度和平均负载分别为

$$l = 1/(1 - p_b), \quad \rho_w = (1 - p_a)/(1 - p_ap_a)$$

(8)

式 (8) 中，分组平均长度反映了在平稳状态下，分组持续到达的程度。因此， l 越大，业务突发程度就越高。

3.1 非突发业务

图 4, 图 5 给出了 3 种交换结构在非突发业务下的分组丢弃性能。图 5 中的“Analytical”是依据式 (7)，只考虑足够多的转换器解决冲突时的理论计算值。

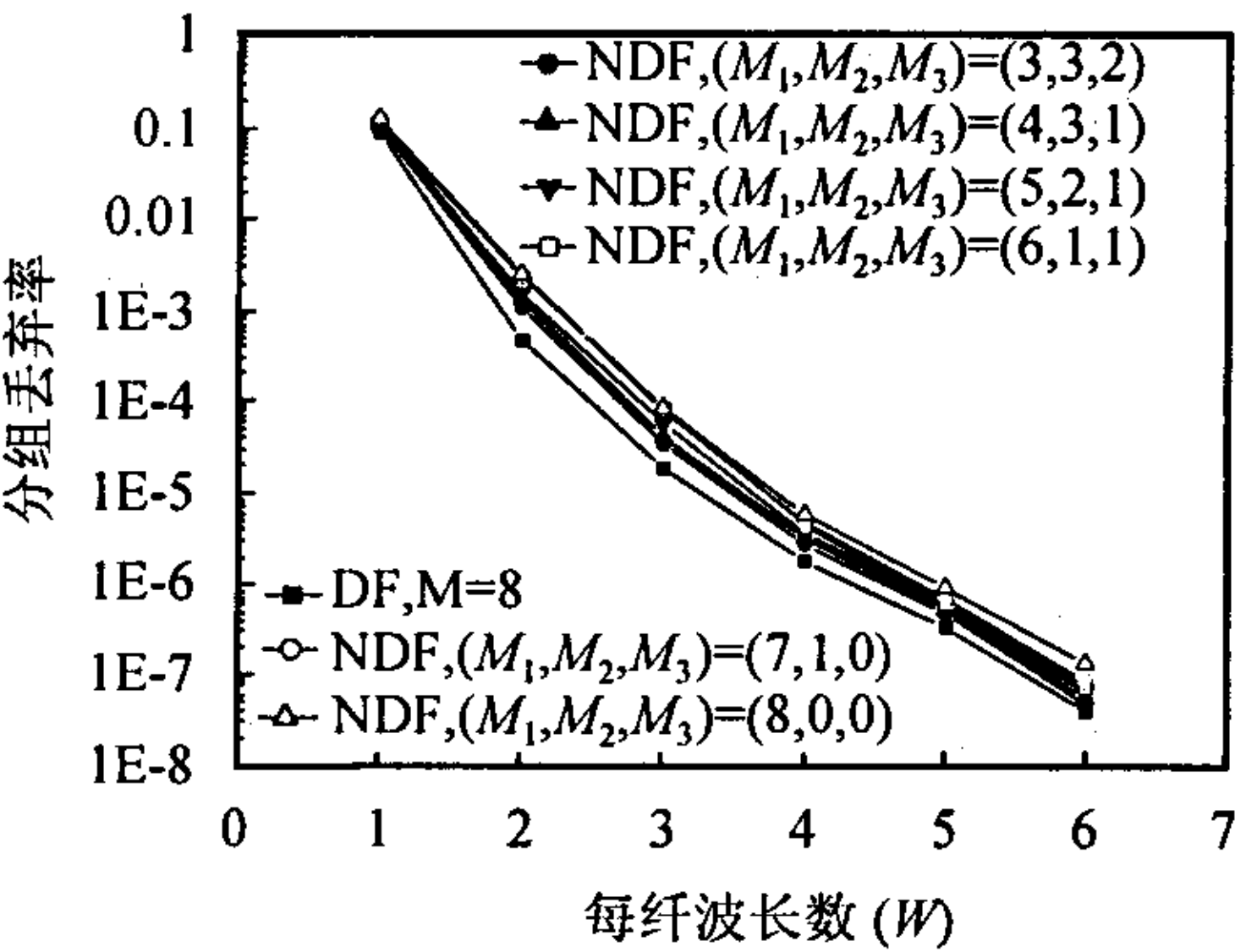


图 4 DF 和 NDF 结构基于非突发业务的分组丢弃率比较

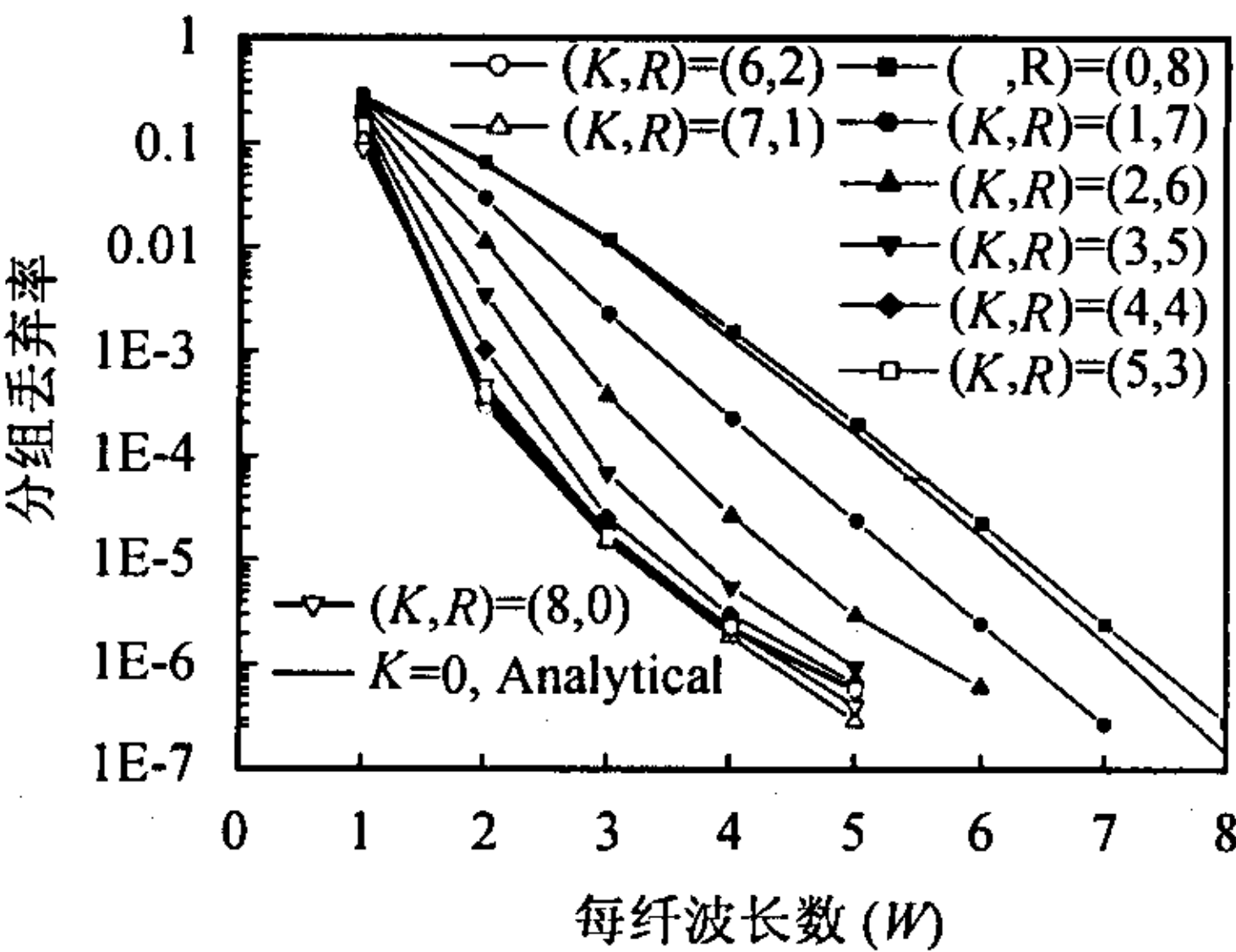


图 5 DFTWC 结构基于非突发业务的分组丢弃性能

从图 4 可以看出，只利用延迟线解决冲突时，对于给定数量的 FDL，光纤总长度越长，则相应的分组丢弃率也愈低。但即使是 NDF 结构中的延迟线取最短长度，即 $(M_1, M_2, M_3) = (8, 0, 0)$ ，它与 DF 结构的分组丢弃率差异也是非常小的。当每纤的波长数不少于 5 个时，分组丢弃率都低于 10^{-6} 。注意该情况下两种结构的延迟线总长度分别为 $8T(\text{NDF})$ 和 $36T(\text{DF})$ 。可见，NDF 结构可以大大减少交换系统的体积。

从图 5 中针对 DFTWC 结构的仿真结果可以看出 4 点：(1) 波长数少于 8 个时，8 个转换器几乎可以逼近理论分析值；(2) 对于 $K + R = 8$ ，随着 FDL 数量的逐渐增多，分组丢弃率将逐渐下降，但延迟线的总数超过 4 个后，改善就很有限了，最理想的配置是取 $(K, R) = (7, 1)$ ；(3) 当每纤波长数不少于 8 个时，不管如何分配转换器和延迟线的个数，分组丢弃率都低于 10^{-6} ；(4) 即使将 NDF 结构配置为 $(M_1, M_2, M_3) = (8, 0, 0)$ ，阻塞性能也比 DFTWC 结构配置为 $(K, R) = (0, 8)$ 时要好。

3.2 突发业务

图 6–图 8 给出了 DF 和 NDF 结构在平均突发长度分别为 $l = 2, 5, 10$ 等 3 种情况下的分组丢弃性能。可见，对于突发业务来说，同样 FDL 数量的 DF 结构比 NDF 结构的分组丢弃率要低得多。这说明，在突发业务环境下，FDL 长度对分组丢弃性能有很大影响。另外，随着分组平均突发长度的增加，分组丢弃率急剧恶化。例如，当 $W = 8$ 时，对于配置为 $(M_1, M_2, M_3) = (3, 3, 2)$ 的 NDF 结构， $l = 2$ 时的分组丢弃率为 5×10^{-5} ，而当 $l = 10$ 时，分组丢弃率急剧恶化为 0.015。从图 6–图 8 还可看出，即使是 DF 结构，表现也不甚理想。

图 9- 图 11 给出了 DFTWC 结构在平均突发长度分别为 $l=2, 5, 12$ 等 3 种情况下的分组丢弃性能。从图 9 可见，对于较低突发度的业务，分组丢弃性能的变化趋势与非突发业务情况十分相似，即 FDL 仍然比转换器更有利于解决冲突。但随着业务平均长度的增加，转换器的作用逐渐显现出来。不妨考察 $W=8$ 时，不同的 FDL 与转换器的配置对分组丢弃率的影响。当 $l=5$ 时，根据图 10， $(K, R) = (5, 3)$ 还能够获得 10^{-6} 的分组丢弃率。当 l 为 12 时，要得到 10^{-6} 的分组丢弃率，转换器的个数必须不少于 6 个，即 $(K, R) = (2, 6)$ 。根据图 9- 图 11 中曲线的变化趋势，可以预见，随着突发程度的进一步增加，必须继续增加转换器的个数或每纤的波长数。

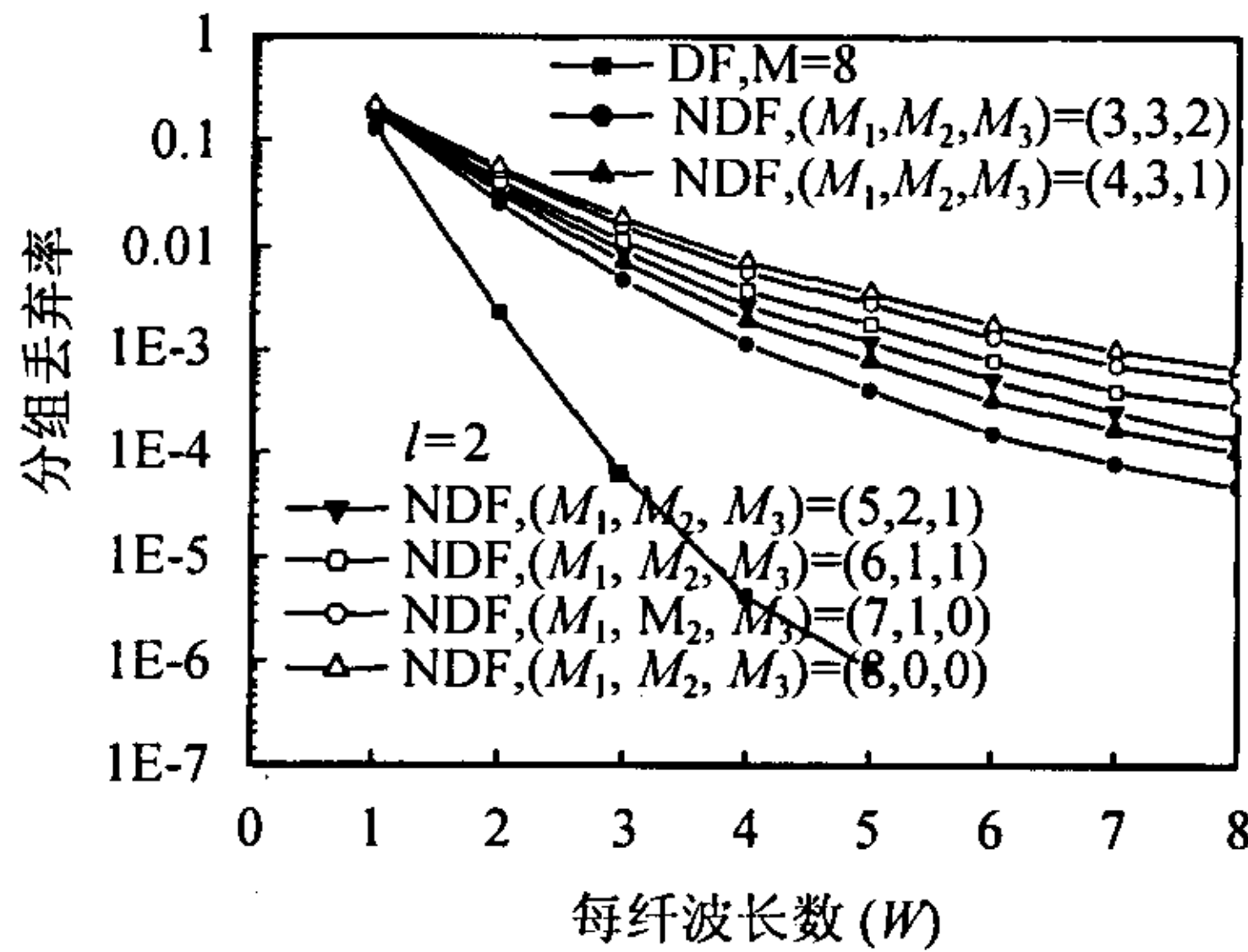


图 6 DF 和 NDF 结构基于突发业务的分组丢弃率比较 ($l = 2$)

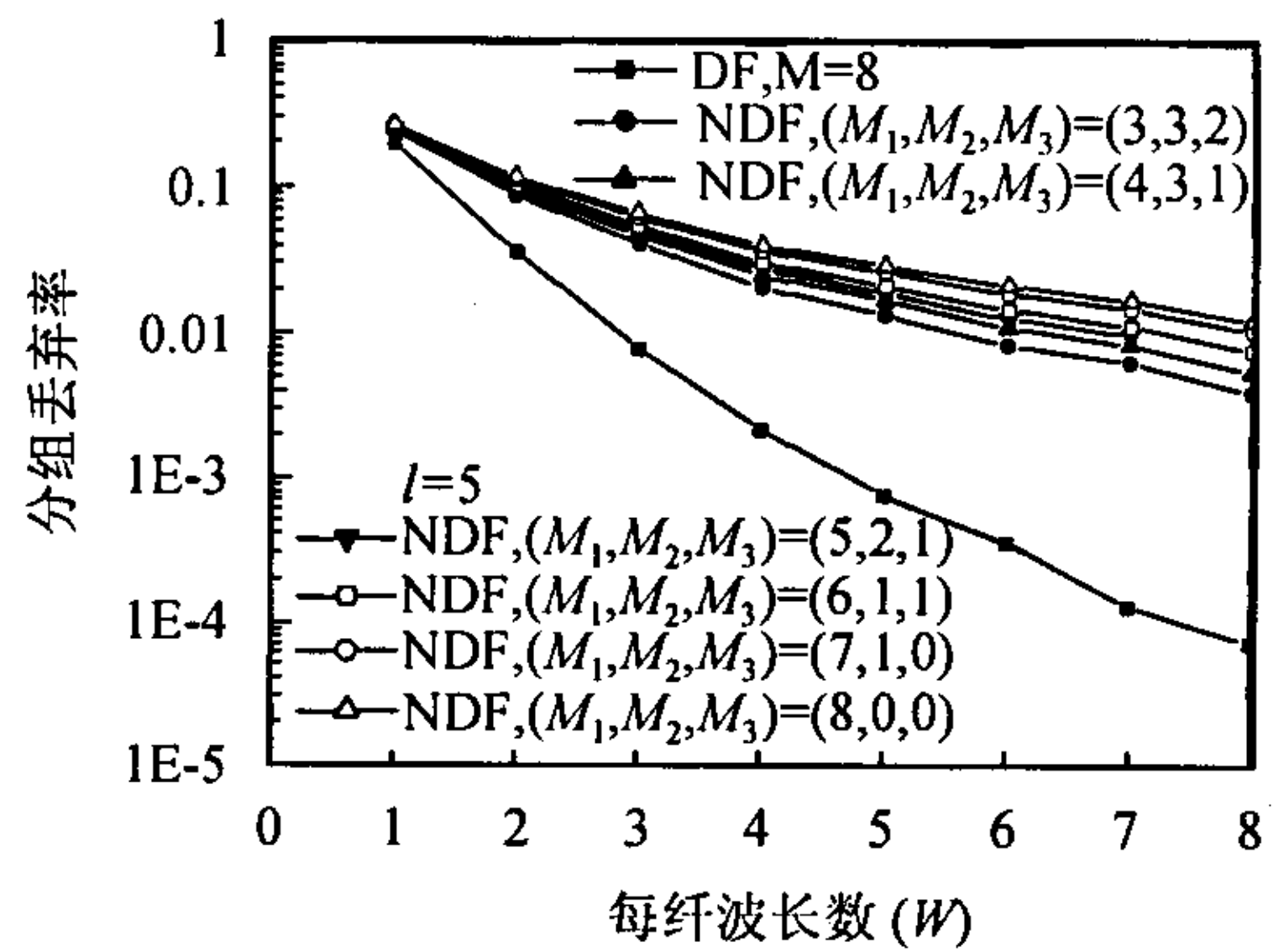


图 7 DF 和 NDF 结构基于突发业务的分组丢弃率比较 ($l = 5$)

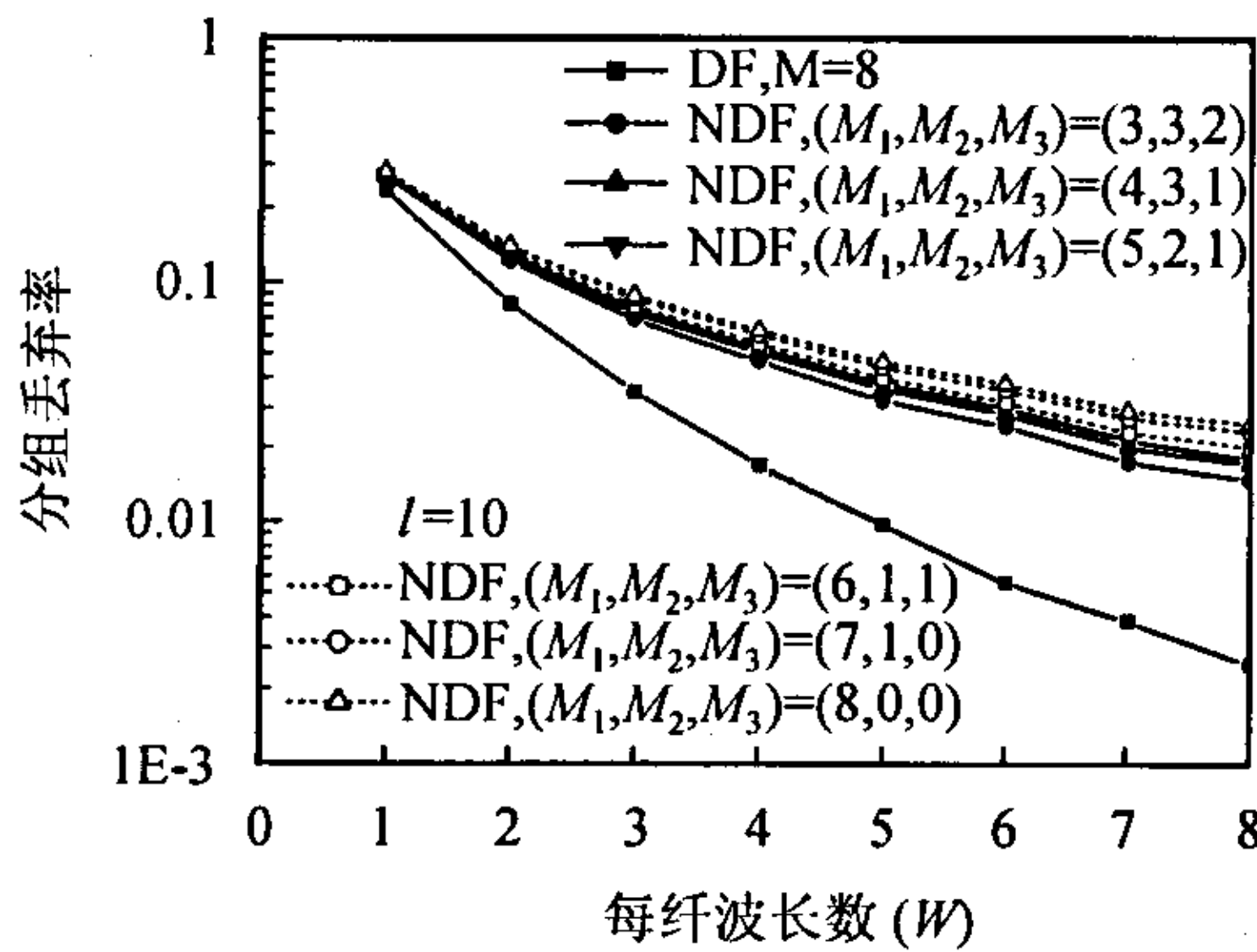


图 8 DF 和 NDF 结构基于突发业务的分组丢弃率比较 ($l = 10$)

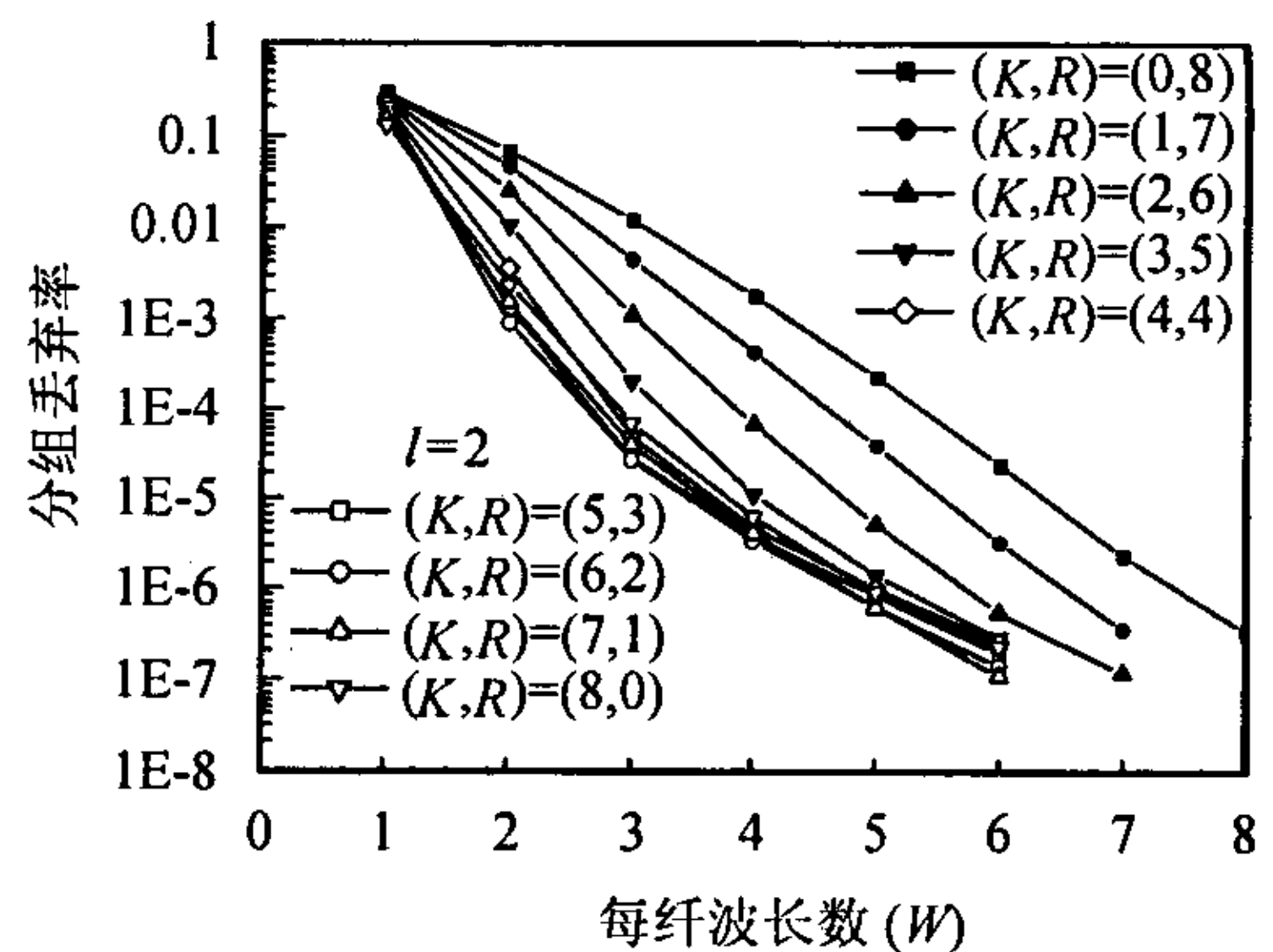


图 9 DFTWC 结构基于突发业务的分组丢弃率 ($l = 2$)

波长转换器比 FDL 对突发业务更加鲁棒，可以这样理解：FDL 通过缓存减少了当前时隙的冲突，但增加了将来时隙的冲突，因此，不适合于高突发性的业务。波长转换器则与之不同，它能够立即处理冲突，并且对随后到来的时隙没有造成任何影响，因此，更适合于突发业务。

由上面的仿真结果及分析，可以得到交换结构设计的以下原则：(1) 应尽量结合目前成熟的 WDM 技术，对负载进行分担，在一定程度上缓解冲突发生的几率；(2) 将 FDL 或波长转换器设计成能为每个输入信道共享的形式，以节约大量的 FDL 及昂贵的波长转换器；(3) 若应用环境是非突发性业务，应尽量采用非简并形式的 FDL 解决冲突，以取得最小的系统体积和成本；(4) 对于中低突发度的业务，可以使用简并形式的 FDL 或将简并形式的 FDL 和少量的波

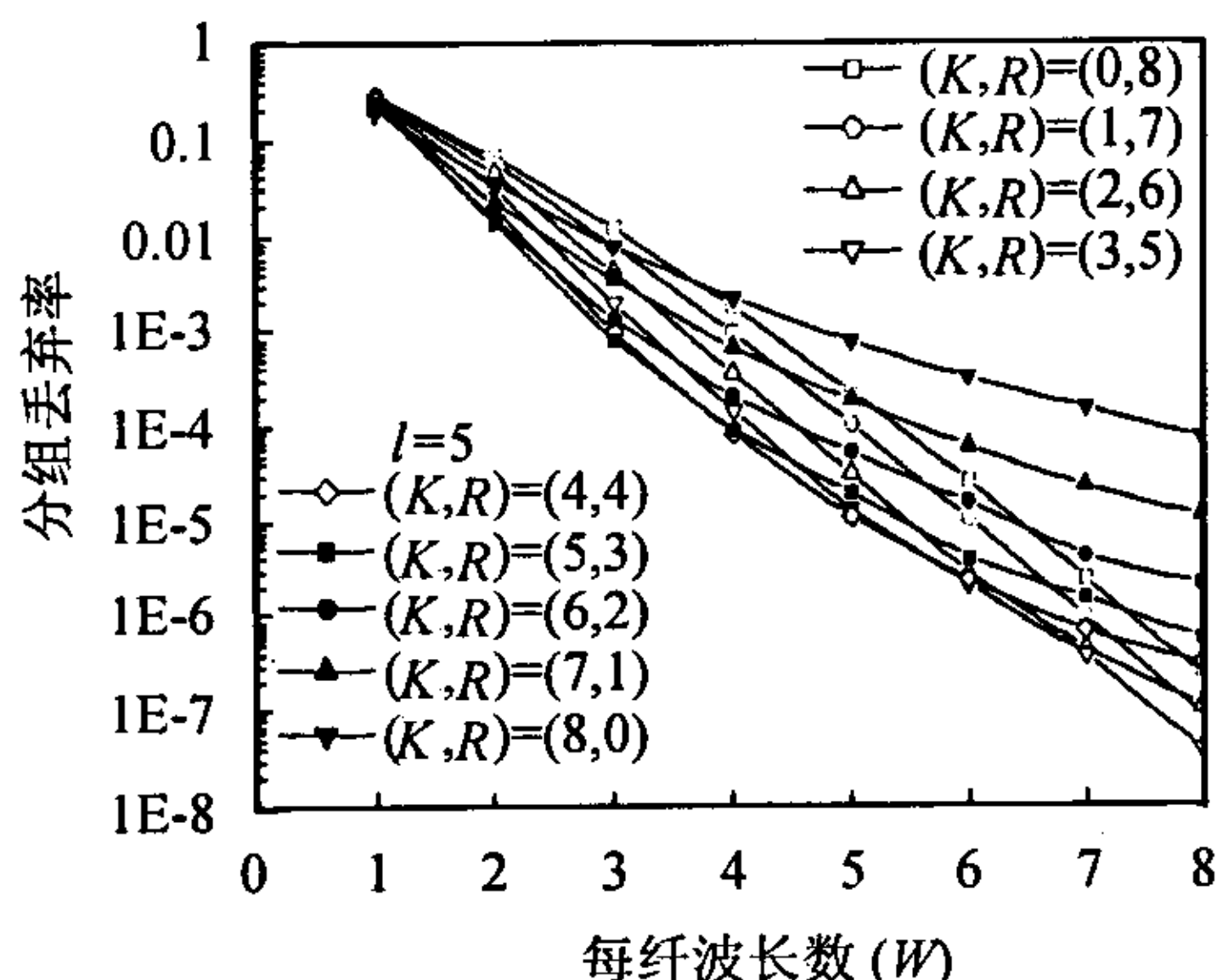


图 10 DFTWC 结构基于突发业务的分组丢弃率 ($l = 5$)

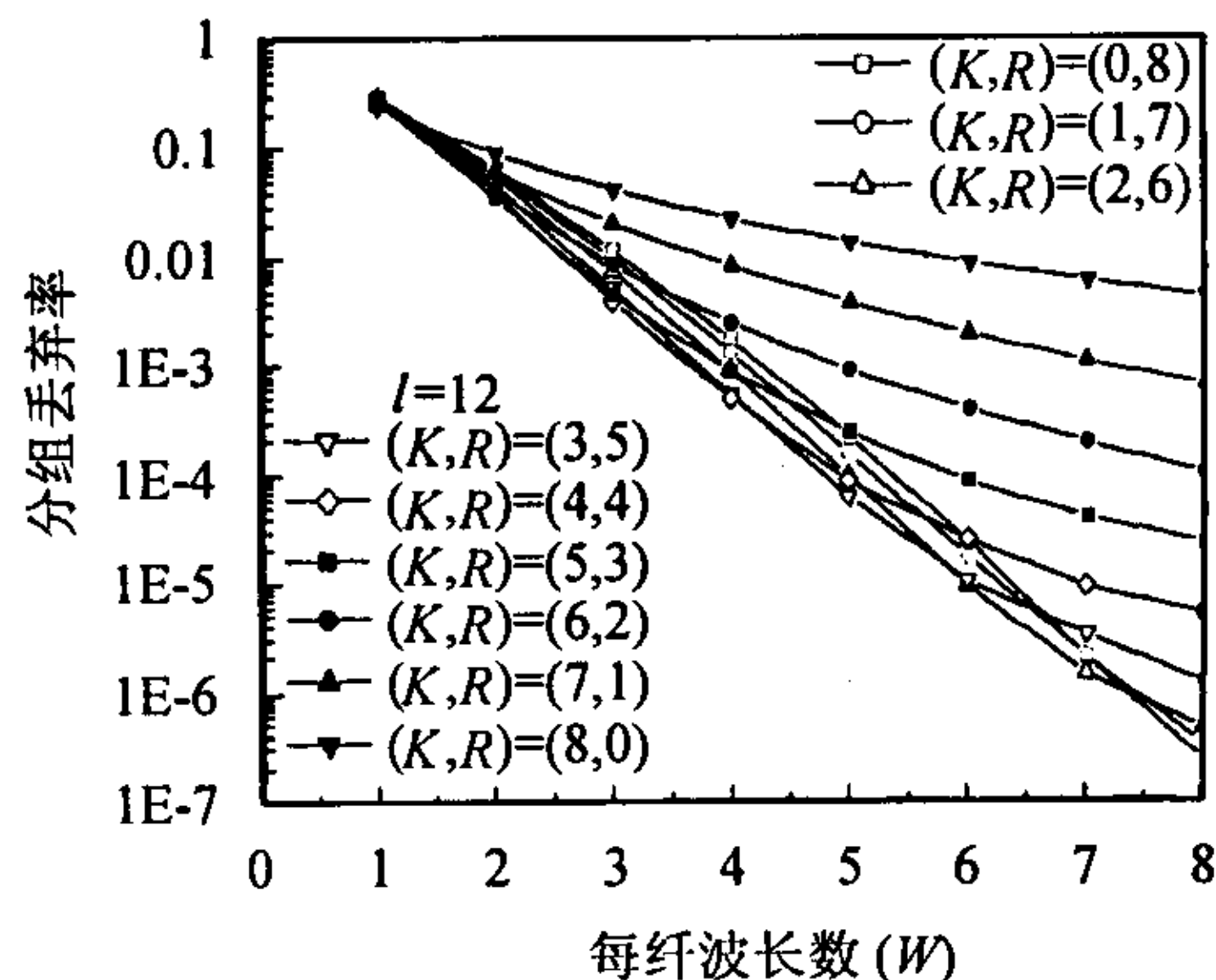


图 11 DFTWC 结构基于突发业务的分组丢弃率 ($l = 12$)

长转换器结合使用,以取得系统体积和成本的折衷;(5)对于高突发度的业务,要满足分组丢弃性能,必须使用更多的波长转换器。系统成本肯定高于前述情况,但无论如何,基于反馈结构形式的、结合使用 FDL 和波长转换器的设计将有助于减少系统的体积及成本。

4 结束语

本文研究了基于时隙的光分组交换网络的冲突处理问题,提出了两种新的,异于传统设计的交换结构。研究表明,对于给定的分组丢弃率要求,本文的两种交换结构能够取得较好的体积和成本折衷,能很好地分别应用于非突发业务和突发业务环境。进一步的研究包括:研究非时隙网络环境下,本文这两种交换结构的性能。

参 考 文 献

- [1] Haas Z. The staggering switch: An electronically controlled optical packet switch [J]. *IEEE J. of Lightwave Technology*, 1993, 11(5/6): 925-936.
- [2] Hunter D K, Chia M C, Andonovic I. Buffering in optical packet switches [J]. *IEEE J. of Lightwave Technology*, 1998, 16(12): 2081-2094.
- [3] Eramo V, Listanti M, Pacifici P. A comparison study on the number of wavelength converters needed in synchronous and asynchronous all-optical switching architectures [J]. *IEEE J. of Lightwave Technology*, 2003, 21(2): 340-355.
- [4] Yao Shun, Mukherjee B, Yoo S J Ben, Dixit S. A unified study of contention-resolution schemes in optical packet-switched networks [J]. *IEEE J. of Lightwave Technology*, 2003, 21(3): 672-683.
- [5] El-Bawab T S, Shin Jong-Dug. Optical packet switching in core networks: between vision and reality [J]. *IEEE Communications Magazine*, 2002, 40(9): 60-65.
- [6] Danielsen S L, et al.. Optical packet switched network layer without optical buffers [J]. *IEEE Photonics Technology Letters*, 1998, 10(6): 896-898.
- [7] Danielsen S L, et al.. Wavelength conversion in optical packet switching [J]. *IEEE J. of Lightwave Technology*, 1998, 16(12): 2095-2108.
- [8] Eramo V, Listanti M. Wavelength converter sharing in a WDM optical packet switch: dimensioning and performance issues [J]. *Computer Networks*, 2000, 32(8): 633-651.
- [9] Zhang Zhizhong, Zeng Qingji. Wavelength converter and fiber delay-line sharing in WDM optical packet switches: Dimensioning and performance issues [C]. *SPIE Opticomm 2003*, TX, Oct., 2003, 5285: 267-277.

张治中: 男, 1972 年生, 博士, 教授, 从事光通信技术的研究与开发工作。
 曾庆济: 男, 1938 年生, 教授, 博士生导师, 从事光通信技术的研究与开发工作。
 王建新: 女, 1971 年生, 博士生, 从事光通信技术的研究与开发工作。
 雒江涛: 男, 1971 年生, 博士, 教授, 从事光通信技术的研究与开发工作。