

输入队列交换机中嵌套周期流优化调度问题的复杂性分析

吴 俊 李 斌

(扬州大学信息工程学院 江苏 扬州 225009)

摘 要 许多网络应用需要网络交换节点能保证分组转发的时延,周期流量的调度是提供这一保证的重要手段.在流量负荷过载的情况下,如何进行优化调度是该领域的重要课题.文中首先依据交换机吞吐率和呼损率两个性能指标,分别定义了两种交换机周期流量调度的最优化问题.为了分析这些优化调度问题的复杂性,文中定义了一种受限的 Max2Sat 问题,并证明该 Max2Sat 问题是 NP 完全的.然后,通过将该问题多项式归约到交换机周期流优化调度问题,证明了仅有 1 和 2 嵌套周期流的交换机优化调度问题是强 NP 完全问题.并进一步利用该结果证明了任意嵌套周期的优化调度问题也是 NP 难的.

关键词 输入队列;交换机;分组调度;周期流;硬实时;NP 完全

中图法分类号 TP393 **DOI 号:** 10.3724/SP.J.1016.2010.00055

On the Complexity of Optimal Scheduling Multi-Rate Nested Periodic Traffic in an Input-Queued Switch

WU Jun LI Bin

(School of Information Engineering, Yangzhou University, Yangzhou, Jiangsu 225009)

Abstract Many applications need that the switching nodes in a network can guarantee the packet deadlines. Multi-rate periodic traffic scheduling is an important method for providing such guarantee. Under overloading traffics, making an optimal scheduling is a key issue. In this paper, two optimal scheduling problems, respectively in terms of switch throughput and call congestion ration, are proposed. In order to analysis the complexity, the authors introduce a restricted Max2Sat problem, and show that the restricted Max2Sat problem is NP complete. Then, a polynomial-time reduction from the Max2Sat problem to the optimal scheduling problem is given for proving that the optimal scheduling problems with only one and two periods are strongly NPC. And this result is generalized to show that any nested periodic traffic optimal scheduling problems are also NP-hard.

Keywords input-queue; switch; packet scheduling; periodic traffic; hard real-time; NPC

1 引 言

由于输出队列结构的交换机制存在严重的可扩展性问题,目前大量的高速交换节点采用了输入队

列结构^[1-5]或 Broadcast-and-Select 光纤交换机制^[6-10].这类交换节点在分组转发时有着共同的限制:同一时刻任意两个被转发的分组不能涉及同一个输入或输出端口(或收发器),否则将产生冲突.因此在作分组调度决策时不仅要考虑分组调度的目标

还要考虑冲突的解析. 另一方面, 随着多媒体信息的广泛使用, 网络中实时流量所占比重日益增多. 这类实时流量一般要求分组在被转发到目的地时满足一定的时限要求. 这要求网络的交换节点在作资源分配决策时能够考虑到流量的这种需求, 即提供一种 QoS 保证的转发服务.

支持 QoS 的交换机调度问题与网络的性能及服务能力密切相关, 因而得到了广泛而深入的研究. 解决这一问题的思路大致可以分为 3 种: 区分服务方式、软实时调度和硬实时调度. 区分服务方式只能提供统计意义上的质量保证, 一般不能提供刚性的时延保证. 文献[3]提出了一种基于博弈的帧调度算法, 该算法在输入流量不过载的情况下可以达到 100% 的吞吐率且能够对不同的流提供不同带宽的服务. 文献[8]中给出 WDM 光纤网络的一种自适应调度算法 MLF-TTS, 该算法兼顾了对实时流量的响应和非实时流量的性能. 但由于这类算法难以提供刚性的服务质量保证, 不能满足对时延、带宽等有着严格要求的应用.

软实时调度算法一般通过对服务带宽的确保来提供时延的保证. 这种带宽的保证可以通过将 PGPS (Packet-by-Packet Generalized Processor Sharing) 方法引入到交换机调度中加以解决. 文献[2]将 PGPS 与稳定婚姻匹配算法相结合, 为每条虚电路提供了服务带宽的保证. 文献[5]采用 Birkhoff-von Neumann 算法将流量矩阵分解再结合 PGPS 来提供每条分组流的带宽保证. PGPS 本质上属于 fluid-track 方法, 文献[4, 11]进一步深入地研究了交换机调度中的 fluid-track 方法. 文献[4]的结果表明采用 fluid-track 类的调度算法, 交换机滞留 (backlogged) 的分组数是 $O(N^2)$. 文献[11]显示最优的 fluid-track 策略只存在于 2×2 交换机. 这表明, 交换机软实时调度算法只能提供较松的时延界保证.

硬实时调度是交换机调度的一个重要课题, 这不仅因为它可以同时提供时延和带宽的保证, 而且目前的光纤交换节点没有大量的光存储器, 而软实时调度却需要较大的分组缓存队列. 文献[6]讨论了逐分组时限保证可调度问题的复杂性, 通过将 TTD 问题^[12] (Time Table Design) 归约到该问题, 证明了仅有 3 个时限类硬实时分组可调度问题是 NP 完全的. 文献[1]研究了逐分组时限保证的分组最优调度问题, 证明了当仅有两个时限类且第 1 个类及总和流量不过载的情况下, 这种硬实时调度问题是多项

式可解的, 并基于两个类的调度算法设计了解决多个时限类的启发式算法. 文献[13]进一步证明了两个时限类时, 若第 1 个类和总和的流量过载的最优调度问题仍然是 NP 难的. 文献[6-7]给出了 EDD (Earliest Due Date)、MLF (Minimum Laxity First) 等启发式算法得到可行解的负载条件. 文献[14]研究了周期流的交换机硬实时调度问题, 提出的 NPS (Nested Period Scheduling) 算法能解决嵌套周期的可调度问题, 并给出了 NPS 用于非嵌套周期时的性能界.

本文着重讨论交换机嵌套周期流优化调度问题的复杂性. 首先我们定义了两种交换机的优化调度问题, 通过将一种受限的 Max2Sat 问题多项式归约到这两个优化问题, 证明了即使是只有 1 和 2 两种周期的情况下, 这类优化调度问题也是强 NP 完全的. 且 1 和 2 周期优化调度问题的 NPC 证明方法可以推广至 1 和 k 周期的情况 (k 为任意大于 1 的整数). 最后我们将 1 和 k 周期的最优调度问题多项式规约到任意嵌套周期的情况, 从而证明任意嵌套周期流的最优调度问题是 NP 难的.

2 问题描述

无阻塞交换机的规模是 $N \times N$, 链路和交换机的工作时间被分成大小相等的时间片, 称为时隙 (timeslot). 传输的分组是等长的, 每个时隙传输一个分组. 交换机是同步工作的, 即每个分组的传输均起始于一个时隙的开始. 交换机在转发分组时, 每个时隙需要有一个配置矩阵用以刻画转发决策, 本文将这种配置矩阵称为调度 s . 显然 s 作为一个调度必须满足 $\sum_{i=1}^N s_{ij} \leq 1$ 和 $\sum_{j=1}^N s_{ij} \leq 1$.

本文主要讨论一种上述交换机中分组流的硬实时 (hard real-time) 调度问题. 这里一条分组流指的是一个分组序列, 该序列中的分组来自同一个输入端口且去往同一个输出端口. 这种分组流有着恒定速率, 即每隔 P 个时隙有一个分组, 我们将 P 称为该分组流的周期. 如果所有的流有着相同的相位, 称为是周期同步的, 下文所提及之周期流均为周期同步的. 硬实时调度要求每个分组必须在该流中下一个分组产生以前被调度, 即该分组流中分组调度的时限 (deadline) 为 P . 我们以流量矩阵 F_k 来表示所有周期为 P_k 的流, 其中矩阵的项 $F_k(i, j)$ 表示来自输入端 i 去往输出端 j 的周期为 P_k 的流数量. 当给

定一组周期分别为 P_1, P_2, \dots, P_k , 流量矩阵为 F_1, F_2, \dots, F_k 的流集合, 我们用 m 表示所有周期的最小公倍数, 即 $m = lcm(P_1, P_2, \dots, P_k)$. 周期流调度的任务就是给出交换机的一组调度 $S = \{s_1, s_2, \dots, s_m\}$ 来确定分组的转发. 如果给定的周期满足 $\forall i < j: lcm(P_i, P_j) = P_j$, 则称这类周期流为嵌套(nested)周期流. MP-TSA 问题(Multiple-Period Time Slot Assignment)指当交换机不过载时(即任何时刻交换机的归一化负荷小于 1), 寻找一组调度使得每个分组能够在时限过期前得到转发. 当讨论这类调度问题的复杂性时, 本文以交换机的规模作为问题的度量, 因此我们仅考虑 $m \leq p(N)$ 的情况($m > p(N)$ 必然没有有效算法), $p(\cdot)$ 为一多项式函数. 目前, 嵌套周期流 MP-TSA 问题已被证明是可以在多项式时间内可调度的^[14].

我们这里考虑比 MP-TSA 问题更一般的情况, 即交换机面临的流量是过载的, 这是网络控制中比 MP-TSA 更为现实的问题. 交换机的输入流量可能是过载的, 导致一些分组在调度时必须丢弃. 由于是硬实时调度, 因此同一条流中的分组或者全部被丢弃或者全部被转发. 在调度时, 必须给定决策的优化目标. 一个自然的优化目标是使得交换机的吞吐率最大化. 为此, 我们定义了 TMP-TSA(Throughput-optimal MP-TSA)问题.

定义 1(TMP-TSA).

实例: 设 $TMP = (P, F, R)$, 其中 $P = \{P_1, P_2, \dots, P_k\}$, $F = \{F_1, F_2, \dots, F_k\}$, R 为一正整数, 设 $m = lcm(P_1, P_2, \dots, P_k)$.

问题: 是否存在调度 S 使得 $\sum_{f \in S} m/P(f) \geq R$, 其中 $f \in S$ 表示流 f 被 S 调度, $P(f)$ 表示流 f 的周期?

上述定义是根据各个流对吞吐率的贡献来确定调度优化目标的, 另一种直观的优化目标是交换机用户的满意度, 即交换机服务的呼损率最小化, 这相当于调度尽可能多的流, 可以用被调度流的数量作为度量, 本文将此优化目标的调度问题称作 SMP-TSA(Satisfaction-optimal MP-TSA)问题.

定义 2(SMP-TSA).

实例: 设 $SMP = (P, F, R)$, 其中 $P = \{P_1, P_2, \dots, P_k\}$, $F = \{F_1, F_2, \dots, F_k\}$, R 为一正整数.

问题: 是否存在调度 S 使得 $\|S\| \geq R$, 其中 $\|S\|$ 表示被 S 调度的流的数量?

$$\begin{array}{ccc}
 F_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & s_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & s'_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 F_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & s_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & s'_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 & s_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & s'_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
 \end{array}$$

(a) 一个(1,3)周期流的流量矩阵

(b) TMP-TSA 问题的解

(c) SMP-TSA 问题的解

图 1

图 1 给出了周期流优化调度的一个简单例子, 共有(1,3)两类周期. 从图 1(a)的流量矩阵显示, 1 周期的流只有 $\langle 3, 2 \rangle$ 一条, 3 周期的流有 $\langle 1, 2 \rangle$ 和 $\langle 2, 2 \rangle$ 两条, 输出端口 2 共需要 5 个时隙才能调度所有的分组, 而 $m=3$, 因此输出端口 2 过载. 该例的调度方案共有两种, 图 1(b)的调度方案 S 丢弃了 3 周期的流, 图 1(c)的调度方案 S' 丢弃了 1 周期的流.

$\sum_{f \in S} m/P(f) = 3$, $\sum_{f \in S'} m/P(f) = 2$, 因此 S 是该例子的 TMP-TSA 问题的解. 然而, $\|S\| = 1$, $\|S'\| = 2$, 该例 SMP-TSA 问题的解却是 S' .

3 NMV-(2,3)-MaxSat 问题及其 NP 完全性

Max2Sat 问题是 NP 完全理论中的一个重要问题, 该问题与许多组合优化问题的复杂性分析有关. Max2Sat 问题的一些受限版本仍然是 NP 难的, 寻找这类受限的 Max2Sat 问题一直是该领域的一个研究方向^[15]. 本文为了分析周期流优化调度问题, 定义了如下受限的 Max2Sat 问题.

定义 3(NMV-(2, 3)-MaxSat).

实例: 设 $E = (C, X, R)$, 其中 $C = \{c_1, c_2, \dots, c_K\}$ 为子句集, $X = \{x_1, x_2, \dots, x_N\}$ 是变量集, R 为正整数. E 满足如下限制: (1) 任一变量最多出现 3 次; (2) 任一子句最多出现两次.

问题: 是否存在变量 X 的真值指派使得 C 中有 R 个以上的子句被满足.

在上面的定义中如果没有任一子句最多出现两次的限制称作 (2, 3)-MaxSat 问题, 文献[15]证明了 (2, 3)-MaxSat 问题是 NPC 问题. 下面我们证明 NMV-(2, 3)-MaxSat 问题也是 NP 完全的.

引理 1. NMV-(2, 3)-MaxSat 问题是 NP 完全的.

证明. MaxSat 属于 NP 类, 显然其子问题 NMV-(2, 3)-MaxSat 也是 NP 问题.

令 $E = (C, X, R)$ 为 $(2, 3)$ -MaxSat 问题的实例, 其中 $C = \{c_1, c_2, \dots, c_K\}$ 为子句集, $X = \{x_1, x_2, \dots, x_N\}$ 是变量集, R 为正整数. 为了证明 NMV- $(2, 3)$ -MaxSat 是 NP 难的, 我们将 $(2, 3)$ -MaxSat 多项式归约到 NMV- $(2, 3)$ -MaxSat.

设集合 $Y = \{y \in X \mid \sum_{i=1}^K |c_i \cap \{y\}| = 3 \vee \sum_{i=1}^K |c_i \cap \bar{y}| = 3\}$. 对 $\forall y \in Y$ 作如下变换. 增加变量 y_1, y_2, y_3 , 将变量 y 的 3 次出现分别替换为 y_1, y_2, y_3 , 记变量 y 出现的 3 个子句的集合为 C_y , 替换后的子句集为 C'_y . 再增加变量 $u_1, u_2, u_3, u_4, u_5, u_6$, 并增加子句集 Z_y :

$$Z_y = \left\{ \begin{array}{l} (y_1, u_1), \quad (y_2, u_2), \quad (y_3, u_3) \\ (\bar{y}_1, u_4), \quad (\bar{y}_2, u_5), \quad (\bar{y}_3, u_6) \\ (\bar{u}_1, \bar{u}_5), \quad (\bar{u}_2, \bar{u}_6), \quad (\bar{u}_3, \bar{u}_4) \end{array} \right\}$$

每个 $u_i (i=1, 2, \dots, 6)$ 将仅在 Z_y 中出现两次. 而 $y_i (i=1, 2, 3)$ 还将在被替换的子句中出现一次, 共出现 3 次, 但都以两种字的形式出现. 因此变换后的实例属于 NMV- $(2, 3)$ -MaxSat:

$$E' = (\cup_{y \in Y} (C'_y \cup Z_y)) \cup C \setminus \cup_{y \in Y} C_y, \\ X \cup U \setminus Y, R + 9 |Y|,$$

其中 U 为新增加的变量集.

这一变换将增加 $9|Y|$ 个变量和 $9|Y|$ 个子句, 显然该变换是多项式的. 下面我们证明若存在真值指派使得 E 有解, 当且仅当存在真值指派使得 E' 有解.

如果存在真值指派使得 E 有 R 个子句被满足, 则只要将与每个 $y \in Y$ 所对应的 $y_1, y_2, y_3, u_4, u_5, u_6$ 指派成同 y 一样的值, u_1, u_2, u_3 指派成与 y 相反的值, 其它 X 中变量的赋值保持不变, 则这使得 E' 中与 E 中对应的子句正好有 R 个被满足, 同时每个 Z_y 中的子句也将被满足, 因此 E' 在这种指派下将有 $R + 9|Y|$ 个子句被满足.

反之, 如果存在真值指派使得 E' 有 $R + 9|Y|$ 个子句被满足. 图 2 考察了 Z_y 在给定 y_1, y_2, y_3 值的同时, 并确保前 6 个子句被满足情况下, Z_y 中不能被满足的子句. 从图 2 可以看出, 只有将 y_1, y_2, y_3 指派成相同值才能使得 Z_y 的 9 个子句全部被满足, 其它的指派将至少有一个子句不能满足. 因此, 如果与 $y \in Y$ 所对应的 y_1, y_2, y_3 在 E' 中被指派了不同值, 我们可以做如下变换: 若有两个变量被指派成了 0, 则将与 y 对应的 9 个变量的值改为图 2 中第 1 种指派, 否则改成图 2 中的第 8 种指派. 由于这种变换只

改变了 $y_i (i=1, 2, 3)$ 中一个变量的值, 最多只影响 C'_y 中一个子句的满足性, 而 Z_y 却全部得到了满足. 因此变换后的指派仍然是 E' 的一个解, 且与 $y \in Y$ 所对应的 y_1, y_2, y_3 被指派成了相同值. 利用 E' 的这个真值指派可以构造 E 的指派: 若 $x \in X/Y$, x 的赋值与 E' 中的相同, 否则将 x 赋值成与其对应变量的值. 由于 E' 中至少有 $R + 9|Y|$ 个子句得到了满足, 而 E' 仅比 E 多了 $9|Y|$ 个子句, 因此 E 在此赋值下将至少有 R 个子句被满足. 证毕.

	Index						Unsat. clause		Index						Unsat. clause
	1	2	3	4	5	6			1	2	3	4	5	6	
y	0	0	0				none	y	1	0	0				(\bar{u}_3, \bar{u}_4)
u	1	1	1	0	0	0		u	-	1	1	1	-	-	
y	0	0	1				(\bar{u}_2, \bar{u}_6)	y	1	0	1				$\bar{u}_2, \bar{u}_6)$
u	1	1	-	-	-	1		u	-	1	-	1	-	1	
y	0	1	0				(\bar{u}_1, \bar{u}_5)	y	1	1	0				$\bar{u}_3, \bar{u}_4)$
u	1	-	1	-	1	-		u	-	-	1	1	1	-	
y	0	1	1				(\bar{u}_1, \bar{u}_5)	y	1	1	1				none
u	1	-	-	1	1			u	0	0	0	1	1	1	

图 2 变量 y_1, y_2, y_3 的赋值及不能被满足的子句

4 嵌套周期流优化调度问题的复杂性

在周期流调度问题中, 可以根据各种不同周期的组合来划分周期流调度问题的子问题, 如上节定义的嵌套周期流. 在嵌套周期流的最优调度问题中, 限定周期只有 1 和 2 两种的子问题记作 $(1, 2)$ -TMP-TSA 和 $(1, 2)$ -SMP-TSA.

利用 NMV- $(2, 3)$ -MaxSat 问题的 NP 完全性可以证明 $(1, 2)$ -SMP-TSA 和 $(1, 2)$ -TMP-TSA 问题也是 NP 难的.

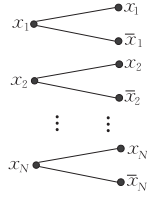
定理 1. $(1, 2)$ -SMP-TSA 问题是 NP 完全的.

证明. 对于 $N \times N$ 交换机, 给定 $(1, 2)$ -SMP-TSA 实例的解, 验证解的正确性需要验证调度的合法性、1 周期的流在 2 个调度中是否有相同的子集和丢弃流的数量是否满足问题的需求. 显然这些验证的时间复杂度都与流的种类成正比, $N \times N$ 的交换机最多有 $O(N^2)$ 种流, 因此验证算法的复杂度是 $O(N^2)$, 这表明 $(1, 2)$ -SMP-TSA 问题是 NP 类的. 下面我们将 NMV- $(2, 3)$ -MaxSat 问题多项式归约到 $(1, 2)$ -SMP-TSA 问题.

令 $E = (C, X, R)$ 为 NMV- $(2, 3)$ -MaxSat 问题的实例, 其中 $C = \{c_1, c_2, \dots, c_K\}$ 为子句集, $X = \{x_1, x_2, \dots, x_N\}$ 是变量集, R 为正整数.

首先, 为变量集 X 创建第 1 类周期的流, 如图 3

(a)所示,每个变量 $x \in X$ 对应一个输入端口 x 、两个输出端口 (x 和 \bar{x}) 和两条流 ($\langle x, x \rangle$ 和 $\langle x, \bar{x} \rangle$), 图 3(b)给出了相应的流量矩阵 F_1^X . 由于第 1 类流的周期是 1, 即只有一个调度, 而每个输入端口关联



(a) 与 X 对应的输入输出端口

$$F_1^X = \begin{matrix} & \begin{matrix} x_1 & \bar{x}_1 & x_2 & \bar{x}_3 & \cdots & x_N & \bar{x}_N \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{matrix} & \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix} \end{matrix}$$

(b) 与 X 对应的第 1 类周期流

图 3 第 1 类周期流构造示意图

然后,为每个子句 $c \in C$ 增加一个输入端口 c 和一个输出端口 c' ; 增加第 2 类周期流 $\langle c, c' \rangle$, 同时若字 $l \in c$ 则增加第 2 类周期流 $\langle c, l \rangle$. 如子句 $c = \langle x, \bar{y} \rangle$ 将对应 $\langle c, c' \rangle, \langle c, x \rangle$ 和 $\langle c, \bar{y} \rangle$ 等 3 条第 2 类周期的流. 由于输出端口 c' 是无冲突的, 我们将 c' 称为哑端口, 与哑端口关联的流称为哑流. 所有哑流的流量矩阵是 $|C|$ 维单位方阵, 记作 $I_{|C|}$. 其它第 2 类流对应流量矩阵是 $|C| \times 2|X|$ 矩阵, 记作 F_2^C .

这样,我们就将 NMV-(2, 3)-MaxSat 归约到了交换机周期流量调度问题, 流量矩阵如图 4 所示, 虽然这里的流量矩阵不是 $N \times N$ 型, 但只需加上 $|X|$ 个空白输入端口即可, 文中不再赘述. 与 E 相应的 (1, 2)-SMP-TSA 问题的实例为 $SMP = (\{1, 2\}, \{F_1, F_2\}, R')$, 其中 $R' = R + |C| + |X|$. 下面我们将证明: 存在真值指派使得 E 有 R 个以上的子句被满足当且仅当 SMP 存在调度 S 满足 $\|S\| \geq R'$.

$$F_1 = \left[\begin{array}{c|c} F_1^X & \mathbf{0}_{|X| \times |C|} \\ \hline \mathbf{0}_{|C| \times |X|} & \mathbf{0}_{|C| \times |C|} \end{array} \right] \quad F_2 = \left[\begin{array}{c|c} \mathbf{0}_{|X| \times |X|} & \mathbf{0}_{|X| \times |C|} \\ \hline F_2^C & I_{|C|} \end{array} \right]$$

图 4 两类周期流的流量矩阵

若存在真值指派使得 E 有 R 个子句被满足, 则构造 SMP 的调度如下:

(1) 设 1 周期类中被调度的流集合为 F_1^S . 对 $\forall c \in C$, 如果 $x = 1$, 则令流 $\langle x, \bar{x} \rangle \in F_1^S$, 否则令流 $\langle x, x \rangle \in F_1^S$.

(2) 设 2 周期类中被调度的流集合为 F_2^S . 对 $\forall c \in C$, 任选一使得 c 成真的字 l , 令 $\langle c, l \rangle \in F_2^S$ 和 $\langle c, c' \rangle \in F_2^S$, 另一条与 c 相关的流丢弃.

由于与任一成真的字 l 相关的 1 周期流被丢弃, 又因为 NMV-(2, 3)-MaxSat 问题的实例中与每个字相关的子句数目不超过 2 个, 因此 l 相应的输出端口的归一化负荷 ≤ 1 . 同时, 容易验证其它端口也不过载. 这样 F_1^S 和 F_2^S 就构成一个负荷不过载且

着两条流, 因此这两条流中至少将有一条被丢弃. 如果与每个变量对应的两条流都正好丢弃一条, 那么每个变量将有一个相应的输出端口空闲, 这些空闲的输出端口将构成一个真值指派.

最大周期为 2 的嵌套周期流, 即存在调度 S 使得 F_1^S 和 F_2^S 中的流在时限超时前被调度. 在上述调度构造的第 1 步共有 $|X|$ 条流被调度, 第 2 步共有 $|C| + |R|$ 条流被调度, 因此 $\|S\| = R'$.

反之, 若 SMP 存在调度 S 使得 $\|S\| \geq R'$. 首先我们将 S 改造成 S' 使得 $\|S'\| \geq \|S\|$ 且任意输入端口 x 在 S' 中的负荷为 1. 不妨设输入端口 x 在调度 S 中的负荷为 0. 由于此 SMP 是由 NMV-(2, 3)-MaxSat 实例变换而来, 因此与输出端口 x 和 \bar{x} 关联的 2 周期流最多 3 条, 故而必有一个输出端口 $o \in \{x, \bar{x}\}$ 在调度 S 中关联的流不超过一条. 我们将 1 周期流 $\langle x, o \rangle$ 替换 S 中关联于 o 的流得到调度 S' , 显然 $\|S'\| \geq \|S\|$ 且输入端口 x 在调度 S' 中的负荷为 1. 这样, 在 S' 中与每个输入端 x 关联的两条流中丢且仅被丢弃了一条, 即 1 周期流在 S' 中被调用了 $|X|$ 条. 根据 R' 的定义, 至少有 $|C| + |R|$ 条 2 周期流得到了调度, 又由于与每个子句对应的哑流只有 1 条, 故至少有 R 个子句的非哑流被调度. 因此如果用 1 周期丢弃流所关联的输出端口作为 E 的真值指派, 将至少有 R 个子句被满足. 证毕.

对于 (1, 2)-TMP-TSA 问题, 我们可以采用定理 1 中类似的方法将 NMV-(2, 3)-MaxSat 的实例多项式归约到 (1, 2)-TMP-TSA 的实例. 因此有如下定理.

定理 2. (1, 2)-TMP-TSA 问题是 NP 完全的.

证明. 给定 $E = (C, X, R)$ 为 NMV-(2, 3)-MaxSat 问题的实例, 按照定理中的步骤构造如图 3 和图 4 所示的周期流量. 所构造的 TMP 实例与定理 1 中的区别在于 R' 的值, 此时我们令 $R' = R + |C| + 2|X|$. 下面我们证明 NMV-(2, 3)-MaxSat 问题的实例有解当且仅当如上构造的 TMP 实例有解.

若 NMV-(2, 3)-MaxSat 问题的实例有解, 我们可以采用与定理 1 中相同的方法从 NMV-(2, 3)-

MaxSat 实例 E 的真值指派构造 TMP 的调度 S . 显然在调度 S 中有对应变量指派的 $|X|$ 条 1 周期流被调度, R 条以上与被满足子句对应的 2 周期流和 $|C|$ 条哑流 (2 周期流) 被调度, 因此 $\sum_{f \in S} m/P(f) \geq 2|X| + R + |C| \geq R'$.

另一方面, 如果存在 TMP 的调度 S 使得 R' 被满足, 同定理 1 先将调度 S 改造成调度 S' , 根据调度的必要条件, 有 $|X|$ 条 1 周期流被调度, 这将对应变量的真值指派, 且 2 周期的哑流只有 $|C|$ 条, 因此至少有与子句对应的 R 条 2 周期流在 S' 中被调度, 故在此真值指派下必定存在 R 以上的子句被满足.

证毕.

定理 1 和定理 2 的证明可以推广至周期是 $(1, k)$ 的情况. 只需将上述证明中变换时添加的哑流变为 $(k-1)\mathbf{I}_{|C|}$ 即可.

推论. $(1, k)$ -TMP-TSA 问题和 $(1, k)$ -SMP-TSA 问题是 NP 难的.

证明. 这里我们仅给出 $(1, k)$ -SMP-TSA 的证明, $(1, k)$ -TMP-TSA 与其类似, 不再赘述. 给定 $E=(C, X, R)$ 为 NMV-(2, 3)-MaxSat 问题的实例, 如图 5 构造 $(1, k)$ -SMP-TSA 实例的流量.

$$\mathbf{F}_1 = \begin{bmatrix} \mathbf{F}_1^X & \mathbf{0}_{|X| \times |C|} \\ \mathbf{0}_{|C| \times 2|X|} & \mathbf{0}_{|C| \times |C|} \end{bmatrix} \quad \mathbf{F}_2 = \begin{bmatrix} \mathbf{0}_{|X| \times 2|X|} & \mathbf{0}_{|X| \times |C|} \\ \mathbf{F}_2^C & (k-1)\mathbf{I}_{|C|} \end{bmatrix}$$

图 5 $(1, k)$ 周期流的流量矩阵

其中 \mathbf{F}_1^X 和 \mathbf{F}_2^C 的定义同定理 1, 不同的是第 2 个周期类中为每个子句 c 添加了 $(k-1)$ 条哑流 $\langle c, c' \rangle$. $(1, k)$ 周期流的调度中有 k 个配置矩阵, 由于哑流不与其它子句的流冲突, 所以 $(k-1)$ 条哑流保证了哑流可以调度在 $(k-1)$ 个配置矩阵中, 从而保证了只要子句 c 中字对应的两条第 2 类周期流若都被调度, 则可以在被调度流总数不变的情况下由哑流替换其中一条. 这样, 余下的讨论类似定理 1, 这里从略.

证毕.

嵌套周期流调度问题按周期的组合可以分成若干种子问题, 如周期是 $(1, 2), (1, 2, 4), (3, 6, 12), \dots$. 并且这些子问题之间并没有包含关系, 因此, $(1, 2)$ 周期流优化调度问题的 NP 完全性并不意味着其它子问题也是 NP 难的. 然而, 直觉上 $(1, k)$ 周期的子问题应该是所有这些子问题中最简单的, 事实也是如此, 我们可以把 $(1, k)$ 周期流的调度问题多项式归约至任意嵌套周期流的情况.

定理 3. 任意嵌套周期流 TMP-TSA 问题是 NP 难的.

证明. 设任意嵌套周期 $P = \{P_1, P_2, \dots, P_n\}$, 其中 $P_2/P_1 = k$. 令 $N \times N$ 交换机的 $(1, k)$ -TMP-TSA 问题的实例 $TMP = (\{1, k\}, \{\mathbf{F}_1, \mathbf{F}_2\}, R)$, 任意的嵌套周期. 我们通过将 TMP 嵌入到一个 $2N \times 2N$ 交换机嵌套周期为 P 的实例 $PTMP = (P, \{\mathbf{F}'_1, \mathbf{F}'_2, \dots, \mathbf{F}'_n\}, R')$, 来证明任意嵌套周期流调度问题的难解性.

在变换时首先增加 N 个输入端口和 N 个输出端口, 将 PTMP 中流量矩阵 $\mathbf{F}'_3, \mathbf{F}'_4, \dots, \mathbf{F}'_k$ 置为 0, 然后将 TMP 的流量矩阵分别嵌入到 \mathbf{F}'_1 和 \mathbf{F}'_2 , 并为第 1 类周期流中每个新增加端口增加 $P_1 - 1$ 条流, 具体流量矩阵如图 6 所示, 其中 \mathbf{I}_N 为 N 维单位矩阵. 最后令 $R' = (R + 2kN(P_1 - 1))P_n/P_2$, 显然这一变换是多项式时间的.

$$\mathbf{F}'_1 = \begin{bmatrix} \mathbf{F}_1 & (P_1 - 1)\mathbf{I}_N \\ (P_1 - 1)\mathbf{I}_N & \mathbf{0}_{N \times N} \end{bmatrix} \quad \mathbf{F}'_2 = \begin{bmatrix} \mathbf{F}_2 & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \mathbf{0}_{N \times N} \end{bmatrix}$$

$$\mathbf{F}'_3 = \dots = \mathbf{F}'_n = \mathbf{0}_{2N \times 2N}$$

图 6 PTMP 的流量矩阵



$$s_D = \begin{bmatrix} \mathbf{0}_{N \times N} & \mathbf{I}_k \\ \mathbf{I}_N & \mathbf{0}_{N \times N} \end{bmatrix}$$

图 7 PTMP 的调度 S'

一方面, 若 TMP 存在调度 $S = \{s_1, s_2, \dots, s_k\}$ 满足 R , 按图 7 构造 PTMP 的调度 S' (由于只有第 1、2 类周期流不为 0, 因此 PTMP 的调度只需给出 $P_2 = kP_1$ 个配置矩阵即可). 不妨设 1、2 周期流分别有 u 、 v 条被 S 调度, 这样 PTMP 中第 1 类周期流中共有 $2N(P_1 - 1) + u$ 条周期流被调度, 第 2 类周期中有 v 条周期流被调度. 因此 S' 的目标函数值为

$$\begin{aligned} & (2N(P_1 - 1) + u)P_n/P_1 + vP_n/P_2 \\ &= (2kN(P_1 - 1) + ku + v)P_n/P_2 \\ &\geq (R + 2kN(P_1 - 1))P_n/P_2 = R'. \end{aligned}$$

另一方面, 若 PTMP 存在调度 S 使得 R' 满足, 由于在变换中每条新加入的流与 \mathbf{F}_1 或 \mathbf{F}_2 中的流只有一个端口的冲突且新增加的端口出没有冲突, 因此不难在保持目标函数值不变的情况下将 S 改造成类似图 7 的调度, 这样即可以从中得到 TMP 的调度, 显然这个调度满足 R .

证毕.

$$\mathbf{F}'_1 = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \mathbf{0}_{N \times N} \end{bmatrix} \quad \mathbf{F}'_2 = \begin{bmatrix} \mathbf{F}_2 & (P_1 - 1)\mathbf{I}_N \\ (P_1 - 1)\mathbf{I}_N & \mathbf{0}_{N \times N} \end{bmatrix}$$

$$\mathbf{F}'_3 = \dots = \mathbf{F}'_n = \mathbf{0}_{2N \times 2N}$$

图 8 PSMP 的流量矩阵

定理 4. 任意嵌套周期流 SMP-TSA 问题是 NP 难的.

证明. 与定理 3 的证明类似, 我们可以把 $(1, k)$ -SMP-TSA 问题的实例按图 8 所示嵌入到任意嵌套周期的实例. 接下来的讨论类似于定理 3, 此处从略.

证毕.

5 结束语

本文以吞吐和呼损为优化目标定义了两种交换机周期流的优化调度问题, 并对其中一类子问题——嵌套周期流优化调度的复杂性进行了研究. 我们首先证明了一种特殊 Max2Sat 问题的 NP 完全性, 以该问题为基础证明了 $(1, 2)$ 周期流优化调度问题是 NPC 问题. 然后利用该结果显示任意嵌套周期流的优化调度问题都是 NP 难的. 实际上, 定理 1 和 2 的证明所涉及的 $(1, 2)$ 周期流优化调度问题的实例中没有任何数值小于交换机规模 N 的某个线性函数, 因此, 根据文献[16]中的定义, $(1, 2)$ 周期流优化调度问题是强 NPC 的. 这表明对于任意嵌套周期流优化调度问题不存在伪多项式算法.

本文未讨论的非嵌套周期流优化调度问题的复杂性, 经典的启发式算法 EDD、MLF 等在周期流优化调度问题下的性能界以及高性能的近似算法设计等将作为我们进一步研究的方向.

参 考 文 献

- [1] Lee Y, Lou J Y, Luo J Z, et al. An Efficient Packet Scheduling Algorithm With Deadline Guarantees for Input-Queued Switches. *IEEE/ACM Transactions on Networking*, 2007, 15(1): 212-225
- [2] Kam A, Siu K. Linear-complexity algorithms for QoS support in input-queued switches with no speedup. *IEEE Journal on Selected Areas in Communications*, 1999, 17(6): 1040-1056
- [3] Bianco A, Giaccone P, Leonardi E, et al. A framework for differential frame-based matching algorithms in input-queued

- switches//Li V O K eds. *Proceedings of the IEEE INFOCOM*. Hong Kong, 2004: 1147-1157
- [4] Tabatabaee V, Georgiadis L, Tassiulas L. QoS provisioning and tracking fluid policies in input queuing switches. *IEEE/ACM Transactions on Networking*, 2001, 9(5): 605-617
- [5] Chang C, Chen W, Huang H. Birkhoff-von Neumann input buffered crossbar switches//Moshe Sidi eds. *Proceedings of the IEEE INFOCOM*. Israel, Tel Aviv, 2000: 1614-1623
- [6] Bonuccelli M A, Clò M C. Scheduling of real-time messages in optical broadcast-and-select networks. *IEEE/ACM Transactions on Networking*, 2001, 9(5): 541-552
- [7] Bonuccelli M A, Clò M C. EDD algorithm performance guarantee for periodic hard-real-time scheduling in distributed systems//Atallah M, Weems C eds. *Proceedings of the IPPS/SPDP 1999*. Washington, 1999: 668-677
- [8] Ma M, Hamdi M. An adaptive scheduling algorithm for differentiated services on WDM optical networks. *Computer Communications*, 2004, 27(1): 857-867
- [9] Wen B, Sivalingam K M. Routing, wavelength and time-slot assignment in time division multiplexed wavelength-routed optical WDM networks//Kermani P eds. *Proceedings of the IEEE Infocom*. New York, 2002: 1442-1450
- [10] Rajalakshmi P, Jhunjhunwala A. Routing wavelength and time-slot reassignment algorithms for TDM based optical WDM networks. *Computer Communications*, 2007, 30(1): 3491-3497
- [11] Shen X et al. Deadline guaranteed packet scheduling for overloaded traffic in input-queued switches. *Theoretical Computer Science*, 2008, 409(3): 477-485(in Press)
- [12] Rosenblum M, Goemans M X, Tarokh V. Universal bounds on buffer size for packetizing fluid policies in input queued, crossbar switches//Li V O K eds. *Proceeding of the IEEE INFOCOM*. Hong Kong, 2004: 1614-1623
- [13] Even S, Itai A, Shamir A. On the complexity of timetable and multicommodity flow problems. *SIAM Journal on Computing*, 1976, 5: 691-703
- [14] Giles J, Hajek B. Scheduling multirate periodic traffic in a packet switch//*Proceedings of the Conference Information Sciences and Systems at John Hopkins University*, 1997
- [15] Raman V et al. A simplified NP-complete MAXSAT problem. *Information Processing Letters*, 1998, 65(1): 1-6
- [16] Garey M R, Johnson D S. *Computers and Intractability: A Guide to NP-Completeness Theory*. San Francisco, CA, 1979: 77-106



WU Jun, born in 1970, Ph. D., associate professor. His research interests include network algorithm, protocol engineering and formal method.

LI Bin, born in 1965, Ph. D., professor. His research interests include Web service, distributed AI.

Background

In modern networks, Input-queued switching fabrics or Broadcast-and-Select optical switching structures are widely used. Both the switching systems have a common property: any two packets to be switched in same timeslot should have different input ports and different output ports, otherwise a conflict occurs. On the other hand, voice and video traffics are widespread in Internet, which appeal switches to provide QoS guaranteed service. Hard real-time packet scheduling in input-queued switches is proposed under the aforementioned background.

In this area, there have some results for deadline guaranteed packet schedulable problem, deadline guaranteed optimal packet scheduling problem and periodic traffic schedulable problem. The deadline guaranteed packet schedulable problem is NPC when deadline classes is more than 2, and EDD algorithm can make a schedule only when the traffic load is under 0.5. The deadline guaranteed optimal packet

scheduling problem is NPC even with two deadline classes. A heuristic algorithm, named FIP, was proposed. The periodic traffic schedulable problem with nested period is polynomially resolvable. While the complexity of the periodic traffic schedulable problem is not clear. And an realistic problem, the periodic traffic optimal scheduling problem, is still empty.

This work proposes two optimal objects for periodic traffic scheduling in IQ switches. And it shows that these two problems are NPC even only with two period classes one and two. This result is interesting because the problems with one and two period seem quite simple. To the best of our knowledge, this result is given firstly. This work is partially supported by Natural Science Foundation of Jiangsu Province under grant No. BK2009698 and BK2007074, and the Natural Science Foundation of No. 60873219 and No. 60903161.