

基于 Web 服务的工作流执行优化方法探讨

胡文江,高永兵,樊瑞民,张 健

(内蒙古科技大学 信息工程学院,内蒙古 包头 014010)

摘要: 工作流优化有多种不同的手段,按照优化操作的对象不同,可以分为结构优化和参数优化。资源优化是工作流参数优化领域的热点,优化资源数量可以最大化地改善工作流时间性能。优化算法主要涉及遗传算法、基于 Petri 网结构的并行优化和与扩展 Petri 网结合的遗传算法等。工作流验证目的是在过程设计时检验工作流的正确性,避免执行时出现异常。在工作流模型实际实施之前,探测其中可能存在的各种过程异常可以降低工作流运行时的停产、检查和修复的成本,具有重大的经济意义。车间作业调度问题是一类最具有般性的生产调度问题,采用这种新型的扩展 Petri 网对调度问题进行建模,结合遗传算法对该调度问题进行优化,最后得到了问题的最优解。这种基于扩展 Petri 网的遗传算法具有较高的通用性。

关键词: 工作流优化;自动机;遗传算法;Petri 网

中图分类号: TP311.5

文献标识码: A

文章编号: 1673-629X(2009)06-0156-04

Research of Workflow Optimal Strategy Based on Web Services

HU Wen-jiang, GAO Yong-bing, FAN Rui-min, ZHANG Jian

(Sch. of Information Engineering, Inner Mongolia Univ. of Science and Tech., Baotou 014010, China)

Abstract: Workflow optimization includes a number of different means. According to the optimization of operating at different targets, workflow optimization consists of structural optimization and quantitative optimization. Resource optimization is a hot spot in the field of the workflow optimization and optimizing the amount of resources can maximize the flow of time so as to improve performance. Optimization related to mainly genetic algorithms, parallel structural optimization with Petri net, and the genetic algorithm based on extended Petri nets (EPN) etc. The workflow verification is mainly testing of the accuracy in the process of designing workflow and it can avoid unusual. Before the workflow model being used in the actual implementation, the verification could detect the abnormal existence, therefore the process be able to reduce the abnormal flow of the cut-off operation, inspection and repair costs, so it is of great economic significance. Job-Shop scheduling problem is the most general of the scheduling problem, the new type of extended Petri nets to model for the scheduling problem, combined with genetic algorithm to optimize the scheduling problem and finally got the optimal solution. The genetic algorithm based expansion Petri net has higher universality than others.

Key words: workflow optimization; DFA; genetic algorithm; Petri nets

0 引言

工作流技术是近年来发展最为迅速的计算机应用技术之一,已被广泛地应用于办公自动化、业务流程重组及其他需要规划和管理业务流程的领域。工作流是针对工作中具有固定程序的常规活动而提出的一个概念。通过将工作活动分解成定义良好的任务、角色、规则和过程来进行和监控,达到提高生产组织水平和工作效率的目的。工作流技术为企业更好地实现经营目标提供了先进的手段。现在工作流管理技术已成功运

用到银行、保险、制造业等各个领域。一个工作流通常需要许多不同的活动对象参与,而各活动对象的工作过程,如调度、等待、协调等,需要消耗一定的人力、时间和数据等资源。在工作流的活动期间合理地分配资源是保证工作流顺畅、合理工作的必要条件,优化资源的分配不但能提高工作流实例的工作效率,而且能提高相关资源的利用率。

1 工作流验证目的意义

工作流优化是指为了达到指定的性能目标而对工作流进行的改进。文中从两个方面(工作流验证和优化方法)讨论工作流执行效率优化问题,一方面是工作流的验证,其目的是在过程设计时检验工作流的正确

收稿日期:2008-09-11;修回日期:2008-12-14

基金项目:内蒙古自然科学基金资助项目(200607010810)

作者简介:胡文江(1959-),男,教授,硕士研究生导师,现从事网络技术和数据库研究。

性,避免执行时出现异常。过程异常包括死锁(deadlock)、死任务(deadtask)、活锁(livelock)等。在工作流模型实际实施之前,探测其中可能存在的各种过程异常可以降低工作流运行时的停产、检查和修复的成本,具有重大的经济意义。

工作流模型是指工作流过程模型,工作流过程模型主要用来描述业务过程,是业务过程的抽象化表示,定义了工作流管理系统执行控制的任务及任务关系。由于工作流建模过程不可避免地会引入错误或存在不合理的结构,必须对工作流模型进行定性和定量分析,以改进模型,实现业务流程优化等。模型的定量分析是指模型的验证和证实,定性分析是指模型的性能分析。

模型验证^[1]就是检验模型结构的静态正确性和运行时的动态正确性。模型证实从一般模型分析的角度来讲仅是证实模型完成其功能的有效性,即确认工作流模型的行为序列是否与要描述的真实世界中的业务流程具有一致性。这种一致性包含两个方面内容:一是证实工作流模型的行为序列集能实现其所模拟的业务流程的全部功能;二是证实工作流模型每个行为序列都对应其所模拟的业务过程的一种执行情况。模型性能分析是对模型的各种性能指标的定量分析,如资源利用率、平均执行时间、平均等待队列长度等。模型验证是模型证实和性能分析的基础。在保证工作流正确性基础之上,如何来进一步提高工作流的执行效率。接下来重点讨论另一方面内容——优化方法策略。

2 优化策略方法

工作流优化有多种不同的手段,根据优化操作的对象不同,可以分为结构优化和参数优化。结构优化是指对工作流的模型的结构进行的优化,如改变事务处理的路径,增加或删除活动等;参数优化则是修改工作流的各类参数,如修改各类资源的数量,修改分配给活动的资源数量等,资源优化是工作流参数优化领域的热点。

在以往的研究中,人们比较关注资源数量的优化分配。如Goldratt和Cox^[2]对于资源数量优化问题,提出了一种基于性能瓶颈的资源分配方法。Van Hee等人^[3]提出了一种边际分配法,并证明了对于某些特殊结构的工作流模型来说,这种分配方法是最优的。但这种方法不适合于包括并发结构的工作流模型^[3]。除了数量上的优化,资源优化还有其专业化/一般化程度的优化。资源在工作流管理系统中通常可以扮演不同的角色,参与不同活动的执行。资源扮演不同的角色,参与执行的不同活动的数量,称之为资源专业化/一般

化的程度。资源的专业化和一般化具有各自的优缺点,例如在资源专业化的情况下,活动处理速度快,质量高,但是资源的利用率较低;在资源一般化的情况下,系统更灵活,资源利用率提高,但是活动处理时间增加,在不同的活动间跳转时,还需要一定的准备时间。由于资源专业化和一般化各具优缺点,因此需要在它们之间取得一个较优的平衡,以最优化工作流的性能。Poysieck和Hennaford^[4]强调了资源一般化的优点,Rupp和Russell^[5]、Seidmann和Sundararajan^[6]讨论了资源一般化和资源专业化两者,但未见提出具体的优化方法。

工作流资源进行优化,也就是使各个活动间的资源需求不产生冲突,并且活动间的资源需求的耦合性较小,使工作流实例中的资源需求达到平衡,从而实现各个活动的资源平均等待时间等目标达到最小。但是,工作流实例的各个活动是相互配合、共同完成任务的,因而各个活动的资源需求有一定的依赖性和先后关系。资源优化是工作流参数优化的一类重要方法。时间是分析工作流性能的重要指标。通过对工作流资源的优化,可以在一定程度改善工作流的时间性能。在成本约束下,优化资源专业化/一般化的程度及资源数量,从而最大化地改善工作流时间性能。

2.1 优化算法

所谓优化算法^[7],其实就是一种搜索过程或规则,它是基于某种思想和机制,通过一定的途径或规则来得到满足用户的问题的解。优化理论与算法是一个重要的数学分支,主要讨论在众多的方案中最优化以及怎样找出最优方案,计算出最优的解。就优化机制与行为而分,目前工程中常用的优化算法主要可分为:经典算法、构造型算法、改进型算法、基于系统动态演化的算法和混合算法等。目前该类算法研究成果分散,缺乏系统化研究,对单个算法的改进和对多个算法组合的相关成果不断出现,这种现象对优化领域的发展无任何积极意义。讨论不同算法在解决组合优化问题时的优化效率。通过将优化算法的流程进行抽象,可以从中抽取出五个关键的部分,包括初始化、计算、判断、更新和输出,将这五个部分进行合理组合,构建了优化算法的统一流程框架;通过对比分析算法流和工作流的特点,提出将算法流映射为工作流的方法,利用工作流理论对算法流进行建模;从评价算法模型在结构上的合理性,通过选取时间性能参数和计算性能参数进行试验比较,表明利用统一算法模型建立的组合算法在计算性能上能够体现原有两个算法的优势。

针对资源专业化/一般化的程度及资源数量两个方面,曾有人对传统的经典遗传算法^[8]做出扩展,提出

了一种基于分层嵌套遗传算法^[9]的资源优化方法。该方法采用 workflow 实例平均响应时间作为 workflow 时间性能的评价指标,选取最优的资源配置,从而最大化地改善 workflow 时间性能。同时利用实例说明了资源在专业化和一般化之间达到均衡的重要性。

2.2 基于 Petri 网的结构并行优化

1962 年,德国科学家 Carl Adam Petri 在其博士论文《用自动机通信》中首次提出了 Petri 网理论。经过 40 多年的发展,经典 Petri 网(Classical Petri Net)已被 Petri 本人和其追随者们以多种不同的方式进行了扩展。W. M. P. van der Aalst 提出把 Petri 网应用于 workflow 建模,并论证了 Petri 网可以表述 workflow 管理联盟规定的各种工作流^[10]。这里以 Petri 网作为建模语言,还因为它具有以下优点:

(1) 由于 Petri 网图形表达方式的自然性和准确性,使得模型看起来直观且易理解;

(2) Petri 具有许多分析工具可以借用,如可达图、标识树、不变量等;

(3) Petri 网具有形式化语义,有强大的数学理论支持,这使得可以对建立好的模型进行仿真、分析和验证;

(4) 独立了提供商, Petri 网不属于某个公司的软件包,不会因为厂商并购而消失。

在结构优化中,要进行服务的并行化分析^[11],每个集成的 Web Service 与一个业务流程是相同的。对于每一个集成的 Web Service 所包含的完整功能来说,工作的执行都是多个 Web Service 相互协作完成的。所以流程常常都包括对其它的外部服务的功能的调用。当这些相互协作的服务在运行的时候,没有同时访问同一个服务对象或者运行数据的话,那么它们之间就是相互独立的,它们的执行结果不会对对方的执行结果造成影响,则它们可以以并行化的方式运行。这个原则是下面研究的主要依据所在。研究 Web Service 是否可以并行化调用,就要重点研究流程中的各个 Web Service 是否是相互独立的。根据实际的需要,必须让在运行期间工作流的运行实例可以进行变化,这样它可以根据服务内部变量的数值决定在运行的时候访问不同的外部对象。

可以把是否可以被并行化的服务调用的特征归纳起来,有以下几点:

1) Web Service 中通常会含有若干个内部变量,而这些内部变量决定了其所在 Web Service 的执行过程,从而决定了某个特定的 Web Service 是否会共享访问外部资源。将某个特定 Web Service 的内部变量集的取值当作该 Web Service 的状态,通过分析在不同状态

下, Web Service 的某个服务函数的执行过程,以判定其是否访问外部的共享资源,就可以确定在某个状态下,该 Web Service 的服务函数是否能够并行化执行。

2) 对于含有内部变量的 Web Service 而言,都会有相应的处理内部变量的函数。通过调用这些函数,就会使该 Web Service 处于不同的状态,因此,可以跟踪这些函数的调用次序,就可以确定 Web Service 在某一时刻所处的状态。

3) 通过上述两点的分析,引入 DFA,用 DFA 中间的状态,表示 Web Service 的状态,用 DFA 的有向变换弧,表示 Web Service 中相应的改变 Web Service 内部变量值的方法。这样,就可以通过 DFA 来分析 Web Service 的并行化执行。

这些特征就给了一个静态发现服务并行调用是否可行的评判标准:可以使用内部变量来决定流程过程的处理,可以把服务实例的内部变量看成是服务组件的状态。

2.3 Petri 网结合遗传算法

Petri 网是一种图形化、数学化的建模工具,为离散事件系统的模型化、性能分析和设计提供了一个统一的环境。Petri 网可以表达许多具有多种约束和逻辑关系实际优化问题,再结合行之有效的最优化算法如遗传算法、启发式算法等,求解优化问题是一个值得研究的领域和方法,具有重要的实际意义。由于 Petri 网的运行状态可以表示为变迁序列(变迁的编码),而遗传算法以决策变量的某种形式编码作为运算对象来进行优化计算,因此两者具有某些关联。

遗传算法是一类可用于复杂系统优化计算的鲁棒搜索算法,用以决策变量的某种形式编码作为运算对象来进行优化计算,特别是对一些无数值概念或很难有数值概念,而只有代码概念的优化问题,编码处理方式更显出其独特的优越性。Petri 网的运行状态可以表示为变迁序列(变迁的编码),因此两者具有某些内在的联系。该算法给出了 Petri 网与遗传算法相结合的优化过程,不同问题其实现方法不同。实际计算时,初始染色体编码随机生成,然后根据染色体编码和加工次序确定资源库所的颜色,最后由算法求出最优调度。

例如, Job-Shop 调度问题是一类最具一般性的生产调度问题,也是运筹学的难题之一,是一典型的 NP 问题。首先根据一类优化问题的特点,在时间 Petri 网(timed Petri net)和着色 Petri 网(colored Petri net)相结合的基础上,提出了一种新型的扩展 Petri 网(EPN)^[12]。采用该扩展 Petri 网对一 Job-Shop 调度问题进行建模,结合遗传算法对该调度问题进行优化,

最后得到了问题的最优解。使用 EPN 进行 Job-Shop 调度问题建模时,为便于理解,将库所细分为状态库所、资源(机器)库所和资源获得库所,变迁则分为工序变迁和资源分配变迁。

对于实际优化问题,在时间 Petri 网和着色 Petri 网的基础上,提出了一种新的扩展 Petri 网 (EPN) 模型,在 EPN 引入了库所颜色,以便于优化决策。并在 EPN 中引入了条件矩阵 Q 用于计算和判断变迁的激发,给出了 Petri 网与遗传算法相结合的优化方法和具体算法。在 Petri 网和遗传算法的基础上,提出了一种基于扩展时间 Petri 网的单亲遗传算法解决 Job-Shop 调度问题的新方法。由于算法中遗传算子是对 Petri 网模型中的元素进行操作,与问题空间的元素无关,因此,与其它调度算法相比,它有较高的通用性。

3 结束语

介绍了 workflow 模型验证和工作流优化方法,在优化算法方面做了探讨,对进一步研究基于 Web 服务方式的工作流执行优化有一定的参考价值。模型验证是模型证实和性能分析的基础。使用新的扩展 Petri 网结合遗传算法优化方法,对典型案例生产车间作业 (Job-Shop) 调度问题进行求解,建立了该 Job-Shop 调度的 EPN 模型,采用遗传算法对这个 Petri 网模型进行优化,体现建模和优化方法的有效性。在实际的企业业务过程可能要复杂得多,在工作流过程优化方面的研究还值得进一步深入。

参考文献:

- [1] 周江波,凌 鸿,胥正川. 基于 Petri 网的工作流优化分析[J]. 中国管理科学, 2005, 13(3): 50-55.
- [2] Goldratt E M, Cox J. The Goal[M]. Aldershot, England: Gower, 1993.
- [3] Reijers H A. Design and Control of Workflow Processes[M]. Berlin: Springer-Verlag, 2003.
- [4] Poyssick G, Hannaford S. Workflow Reengineering[M]. Mountain View, California: Adobe Press, 1996.
- [5] Rupp R O, Russell J R. The Golden Rules of Process Redesign[J]. Quality Progress, 1994, 27(12): 85-92.
- [6] Seidmann A, Sundararajan A. The Effects of Task and Information Asymmetry on Business Process Redesign[J]. International Journal of Production Economics, 1997, 50(2-3): 117-128.
- [7] 时亚娟. 面向工作流的优化问题求解方法研究[D]. 天津: 河北工业大学, 2006.
- [8] 王小平, 曹力明. 遗传算法-理论、应用与软件实现[M]. 西安: 西安交通大学出版社, 2002: 10-12.
- [9] 肖志娇, 常会友, 衣 杨. 成本约束下工作流时间最小化的资源配置优化[J]. 系统仿真学报, 2006, 18(11): 3320-3323.
- [10] van der Aalst W M P. The application of Petri nets to Workflow Management[J]. The Journal of Circuits, Systems, and Computers, 1998, 8(1): 21-66.
- [11] 朱庆生, 付 新. 基于 Web Service 的工作流系统及优化[D]. 重庆: 重庆大学, 2005.
- [12] 周卫东, 杨加敏, 贾 磊, 等. 一种 Petri 网结合遗传算法的优化方法及应用[J]. 山东大学学报, 2005, 35(4): 59-67.

(上接第 155 页)

法降低。

参考文献:

- [1] Amilhastre J, Fargier H, Marquis P. Consistency restoration and explanations in dynamic CSPs - Application to configuration[J]. Artificial Intelligence, 2002, 135: 199-234.
- [2] 邵伟平, 刘永贤, 郝永平, 等. 基于分布式约束满足的产品配置研究[J]. 东北大学学报: 自然科学版, 2007, 28(1): 103-106.
- [3] 孙吉贵, 景沈艳. 非二元约束满足问题求解[J]. 计算机学报, 2003, 26(12): 1746-1752.
- [4] 董甲东, 郑春香. 分布式系统的时间同步容错机制研究[J]. 计算机技术与发展, 2008, 18(3): 99-101.
- [5] 王秦辉, 陈恩红, 王煦法. 分布式约束满足问题研究及其进展[J]. 软件学报, 2006, 17(10): 2029-2039.
- [6] Yokoo M, Hirayama K. Algorithms for Distributed Constraint

Satisfaction: A Review[J]. Autonomous Agents and Multi-Agent Systems, 2000(3): 185-207.

- [7] Zivan R, Meisels A. Concurrent search for distributed CSPs[J]. Artificial Intelligence, 2006, 170: 440-461.
- [8] 孙吉贵, 高 健, 张永刚. 一个基于最小冲突修补的动态约束满足求解算法[J]. 计算机研究与发展, 2007, 44(12): 2078-2084.
- [9] Yokoo M, Durfee E H, Ishida T, et al. Distributed constraint satisfaction for formalizing distributed problem solving[C]//in Proc. 12th IEEE Int. Conf. Distributed Comput. Syst. Los Alamitos: IEEE Computer Society Press, 1992: 614-621.
- [10] Faltings B, Yokoo M. Introduction: Special Issue on Distributed Constraint Satisfaction[J]. Artificial Intelligence, 2005, 161: 1-5.
- [11] Gent I P, MacIntyre E, Prosser P, et al. Random constraint satisfaction: flaws and structure[J]. Constraints, 2001, 6(4): 345-372.