

程序挖掘的体系结构研究^{*}

陈章^{1,2} 陈志刚¹

(中南大学信息科学与工程学院 长沙410083)¹ (佛山科学技术学院计算机系 佛山528000)²

摘要 本文在描述程序挖掘的概念基础上,提出了用体系结构的思想来研究程序挖掘过程,并给出了一种三层程序挖掘体系结构实例,增强了整个系统的可扩展性和可维护性,为开发一个更灵活、可伸缩、高效的程序挖掘系统提供了广泛的前景。

关键词 程序挖掘,体系结构,Agent 群体

Research of Program Mining Architecture

CHEN Zhang^{1,2} CHEN Zhi-Gang¹

(College of Information Science and Engineering, Central South University, Changsha 410083)¹

(Department of Computer, Foshan University, Foshan 528000)²

Abstract This paper points out research of program mining with architecture based on describing the concept of program mining. A three-tier architecture of program mining is proposed, which promises to develop the system of program mining.

Keywords Program mining, Architecture, Agents

1 引言

随着 Internet 的飞速发展,人们不仅在其上发布和共享各种信息,而且开始在 Internet 上运行各种应用,为用户提供各种各样的服务,如在线游戏、IP 电话、网上购物、电子商务等等,Internet 已经逐渐从信息发布与共享媒体演变为一个大规模的分布式服务提供平台。随着 Web 普及以及 Web/Server 软件结构的成熟,传统软件有向 Web 环境移植的巨大压力,软件与服务概念互相渗透。在网络环境下基于构件的程序挖掘^[1]已经成为软件工程新的研究方向,被视为提高软件开发效率和质量、实现软件复用的重要方式。

程序挖掘基本思想是利用多个 Agent,分析用户计算请求,从大量的构件资源库中识别、发现、获取所需的构件,并进行自动组装,形成实现用户所需计算功能的程序。通过程序挖掘,可以根据用户的要求,自动从网络构件库中发现组装完成相应计算功能的程序,并在客户端链接执行,完成要求的处理功能,在分布式网络中实现按需计算,为用户提供最大限度的灵活性和定制服务。程序挖掘也是在 Internet 上进行服务管理的重要支撑技术。

在程序挖掘中,特别强调了 Agent 的应用,其原因是近年来 Agent 研究正在和软构件技术以及网络技术相结合。Agent 技术是一种全新的分布式计算技术。一方面,提供了一种全新的系统设计思想,提高了系统内实体的能力,使之比以前具有更强的能力;另一方面,为解决分布式问题提供了有效的途径,为系统中实体的协作提供了有效的概念模型^[2]。Agent 通常包含两个含义:第一,Agent 较普通构件具有自主性,能够对周围环境具有一定的感知性和反应能力;第二,Agent 可以理解为移动构件,它可以根据需要在网络节点间

移动^[3]。

接下来文章简要介绍了程序挖掘的概念示意图和主要包含的内容,在此基础上阐明了研究程序挖掘体系结构的意义,并给出了程序挖掘体系结构的定义。文章第三部分给出了程序挖掘的一种三层体系结构实例,详细说明了其工作机理,各部分之间的关系。最后对全文进行总结,并探讨将来的研究方向。

2 程序挖掘体系结构概念

在文[4]中提出了程序挖掘的概念示意图,见图1。

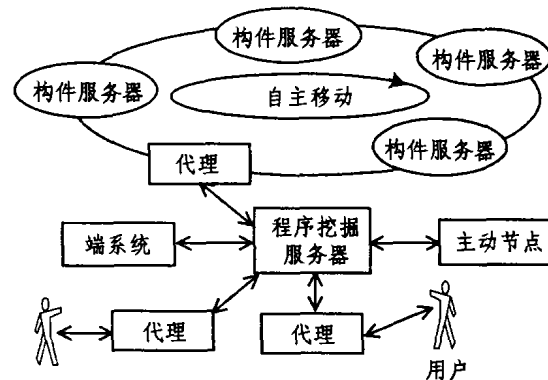


图1 程序挖掘概念示意图

用户的计算请求经智能代理提交到程序挖掘服务器,基于对用户请求的分析,该服务器启动相应的代理在多个构件库之间自主移动,寻找满足要求的构件,然后通过挖掘算法分析构件之间的依赖和调用关系,找出能够实现特定计算功能的构件组合,并将构件组装成满足要求的可执行程序,再通过

^{*} 本课题得到湖南省自然科学基金(02JJY2097)资助。陈章 博士生,副教授,主要研究领域为构件技术、构件库、程序挖掘。陈志刚 教授,博士生导师,主要研究领域为软件工程、网络计算。

代理提交给用户。文[5]中指出程序挖掘过程的六个阶段,分别是:(1)定义统一的构件描述语言;(2)用户需求的描述与分解;(3)构件搜索与获取;(4)构件分析选择;(5)构件组装;(6)程序校验与执行。

文[4]中虽然提出了程序挖掘的概念,但没有从体系结构的角度讨论程序挖掘。从程序挖掘的概念出现至今,以系统的体系结构观念来研究还实属鲜见。我们认为程序挖掘属于复杂的分布式软件系统,文[6]中指出研究软件体系结构(Software Architecture)将会在软件开发中发挥十分重要的作用。它将大型软件系统的总体结构作为研究的对象,认为系统中的计算元素和它们之间交互的高层组织是系统设计的一个关键方面。其研究和实践旨在将一个系统的体系结构显式化,以在高抽象层次处理诸如全局组织和控制结构、功能到计算元素的分配、计算元素间的高层交互等设计问题^[7]。同时,软件过程和软件体系结构是现代软件开发中最关注的两大问题,它们是软件“时空”的外在表现,体系结构影响软件过程制定,而软件过程生成体系结构,一个很好定义的开发过程也是重用的关键^[8]。

另外,由于程序挖掘是基于 Internet 的一个软件服务,Internet 中设备在不断更新换代,技术在不断地提高,网络结构在不断地变化。无论网络的设备、技术和拓扑结构如何变化,对程序挖掘来说其最基本的体系结构应该是不变的,不应当在网络发生新的变化时,就把原有的程序挖掘体系结构推倒重来,这种方法不可取,也是不现实的。因此研究程序挖掘体系结构具有重要的意义。

根据 ODP (开放分布式处理, Open Distributed Processing)关于体系结构的概念,我们给出程序挖掘体系结构的概念。

定义1(程序挖掘体系结构) 定义系统的结构及系统成员间相互关系的一套规则。

在这套规则中提供了对程序挖掘一般的理解和通信;关于如何以及何时使用某个概念的指南,以及有关其属性的信息;根据这些指南或实际的技术,何时可以实现和使用这些概念的指示。

从体系结构的角度来讲,体系结构有多种视图。通常分为概念、逻辑和物理视图。“概念视图”是最抽象的视图,一般用系统用户(非 IT 专业用户)熟悉的术语来描述。概念视图用于定义功能需求和商业用户视图,以便生成业务模型。“逻辑视图”显示了主要的功能组件和它们在系统中的关系,而不涉及功能的实现细节。“物理视图”是最不抽象的,它们表示特定的实现组件和它们之间的关系。物理视图中的每个元素一般都由设计和开发过程来实现,如软件和硬件系统。这里我们谈的程序挖掘体系结构指的是其逻辑视图。

3 程序挖掘模型实例

程序挖掘是一个复杂的分布式过程,文[2]中指出面向 Agent 方法适合开发复杂的分布式系统,它将分布式系统分解成为相对自主的 Agent。每一个 Agent 由一组信念、希望和驱动,多个自主性 Agent 能够以多种灵活的方式实现目标集。本文提出了程序挖掘过程的一种可能模型,为系统地研究程序挖掘做出了一种尝试。

3.1 体系结构

本文提出的是一种三层结构的协同程序挖掘开发模型,这种方案将系统中各种各样的组件划分为“三层”服务,它们

共同组成一个应用服务程序。底层是应用层,中间层是虚拟容器和 Agent 群体,上层是资源层(见图2)。

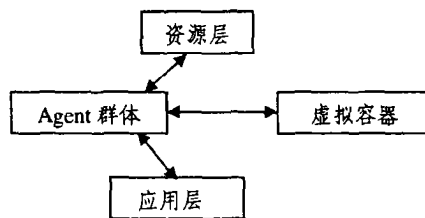


图2 三层程序挖掘体系结构

资源层 这一层的作用是对单个构件库实施控制,与可用构件进行安全握手、监测构件运行状况、统计与付费有关的资源使用数据。

应用层 这层是用户的应用程序。应用程序通过其 API 向中间层发出软件开发服务请求,以交互方式向中间层描述其需求要求并从中间层获得软件服务。

虚拟容器 虚拟容器主要包括容器和容器管理两部分。容器管理负责将资源层提交的受控构件汇集在一起管理;容器(container)暂时存放构件。

Agent 群体采用的是一种树型层次结构的组织方式,根据各 Agent 所负担的职能和所在的层次的不同分为:服务管理 Agent,辅助 Agent 和服务 Agent 三大类,见图3。字典存储着与 Agent 群体相关的知识,如对应不同用户模型的一系列分解方案、不同用户模型下已掌握构件列表及位置等。

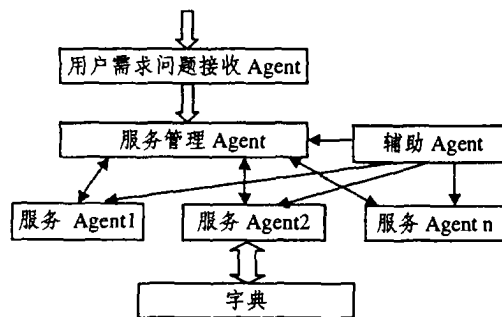


图3 Agent 群体层次结构

3.2 工作机理

(1)应用程序在应用层向中间层的 Agent 群体发出软件服务请求,Agent 群体接受请求并询问需求要求,Agent 群体通过交互方式从应用程序得到需求要求描述。

(2)Agent 群体根据需求要求描述从字典中检索出匹配的用户模型方案和与该模型相关的熟知构件序列,若检索不成功则根据特征值创建新用户模型。

(3)在熟知构件序列基础上分解需求要求(采用混合的方式)得出待搜索的构件序列。

(4)将熟知构件、搜索到的构件放在虚拟容器中筛选、修改、适配,在虚拟容器中设计新构件。

(5)在虚拟容器中组装构件,测试所组装的软件,将测试的结果通知应用层的应用程序。

(6)应用程序审查测试结果,不满意则把结果返回给 Agent 群体。

(7)重复(3)~(6)步骤直至满意为止。

3.3 Agent 群体工作原理

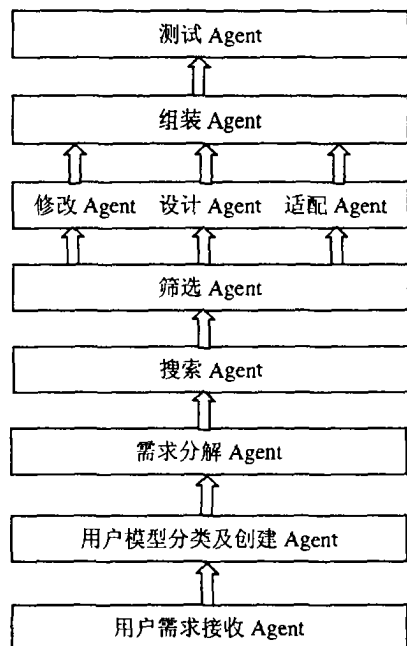


图4 Agent 群体工作流程

如图4所示 Agent 群体工作流程,具体解释如下:

(1)用户需求接收 Agent 负责和应用程序交互,接受问题、分析问题和提取问题特征值如领域特征等。

(2)用户模型分类 Agent 负责从字典中搜索与问题特征值相匹配的用户模型,把与该用户模型相关的知识映像到共享区,如该种模型所熟知的构件序列、构件之间的连接关系、构件所处的构件服务器等等。

(3)用户模型创建 Agent 负责根据问题特征值创建新的用户模型并初始化与其相关的知识。

(4)需求分解 Agent,采用的是一种自上而下和自下而上相混合的方法来分解任务。充分利用已有的领域和构件的知识,把所熟知的构件组合成若干个较大的功能模块,在这些较大的功能模块基础上从上而下地分解服务,得出若干个分解方案列表。

(5)搜索 Agent 负责从方案中提取所需构件的描述,形成构件查询条件,该条件由构件的应用领域、功能描述、类别、接口特征等组合而成。根据构件查询任务,首先在所熟知的构件序列表中查找,然后才访问构件服务器。

(6)筛选 Agent 负责对实现功能相同的、在组合过程中对外部条件有不同的要求的多个构件进行筛选和优化并确定构件组装方案。

(7)修改 Agent,构件有被修改后复用的可能。构件既可以原样复用,也可能需要修改。这是由于通信协议的改变,或者系统功能的扩充和协作构件的升级等原因。为了满足复用者的需求,适当的修改也是必要的。修改 Agent 负责修改构件。

(8)适配 Agent,由于每个构件都为不同的要求而写的,都是基于不同的假设前提下,在新的系统中构件必须适配。适配 Agent 负责选取适配方案,使得构件间的冲突最小。

(9)组装 Agent 负责根据既定的方案将构件粘接在一起。

(10)测试 Agent 负责测试整个服务以保证其满足用户要求所规定的功能、质量和性能等方面的特性。

3.4 体系结构实例的特点

层次结构的程序挖掘模型实例把程序挖掘过程划分为三

个逻辑层,应用层可以把注意力集中在各种用户需求描述表示工具及与用户交互上,不需要考虑互联网中千差万别的差异,不需要知道构件是如何存储和分布的,不需要知道构件是否存在和构件规范。中间层采用的是树型层次结构的组织方式,实现构件挖掘及基于构件的软件开发功能。管理 Agent 负责协调各个服务 Agent,以管理 Agent 为中介实现了服务 Agent 之间的信息交互,避免了服务 Agent 的替换或删除都有可能影响到系统的其他 Agent,使得各个服务 Agent 之间的依赖性及控制相关性较小,提高了系统的可扩展性和可维护性。辅助 Agent 负责协助管理 Agent 工作,它使得程序挖掘过程能够在各种异构平台之上实现。为异构平台构筑了一层通用的、与平台无关的信息和服务交换设施。从而屏蔽了互联网中千差万别的差异,使信息和服务畅通无阻地在计算机之间流动。随着程序挖掘过程实现的次数增加,系统的领域知识和构件知识也相应地增多,如何充分地利用这些知识是基于构件的软件工程所关注的问题。字典存放着 Agent 群体共享的知识如领域知识和构件知识,引入字典为基于软件体系结构的构件化开发提供了客观条件。该开发方法弥补了传统基于构件开发和基于领域开发的不足,在更高层次上实现了软件重用,大大提高了软件的开发效率和开发质量。资源层只提供构件库的管理与控制,为中间层创建构件实例。

总结 本文研究了程序挖掘的概念,阐明了研究程序挖掘体系结构的意义,提出了程序挖掘体系结构的概念并给出了一种程序挖掘过程三层体系结构实例。这种分层体系结构分解了程序挖掘过程,提高了子层的独立性,逻辑结构清晰、简明,实现了动态地配置系统服务 Agent 的功能。提高了灵活性和开发效率。有利于团队的分工合作,可灵活地选择不同层、不同组件工具。为开发一个更灵活、可伸缩、高效的程序挖掘系统提供了广泛的前景。

基于体系结构开发应用系统,能够快速实例化出适应性很强的产品,在功能、性能上满足用户的需求。但是,本文只从逻辑上架构了程序挖掘体系结构的基本框架,重点描述中间层的各组件间的关系。用面向 Agent 方法研究 Agent 群体的工作是今后进一步研究的重点。

参考文献

- 1 张尧学,魏子忠,许克刚,李兴,等. 网络环境下程序挖掘方法及系统. 国家发明专利,专利申请号:0111 8354. 3
- 2 Jennings N R. An Agent-based Approach for Building Complex Software Systems. *Communications of the ACM*, 2001, 44(4)
- 3 Wooldridge M, Jennings N R. *Intelligent agents: Theory and practice*. *The Knowledge Engineering Review*, 1995, 10(2)
- 4 魏子忠. 网络环境下程序挖掘方法的研究:[博士学位论文]. 北京:清华大学, 2001
- 5 樊郁宏,陈松乔. 程序挖掘中需求描述的研究[J]. *计算机工程与应用*, 2002. 10
- 6 Garlan D, Shaw M. *An Introduction to Software Architecture*. In: V Ambriola, G Tortora, eds. *Advances in Software Engineering and Knowledge Engineering Volume 1*. New Jersey: World Scientific Publishing Co, 1993
- 7 Allen R, Garlan D. *Formalizing Architectural Connection*. In: *Proc. of the 16th Intl. Conf. Software Engineering*. IEEE Computer Society, 1994. 1~80
- 8 Perry D E. *Software Architecture and Software Engineering*. 16th World Computer Congress, Beijing, Aug. 2000