



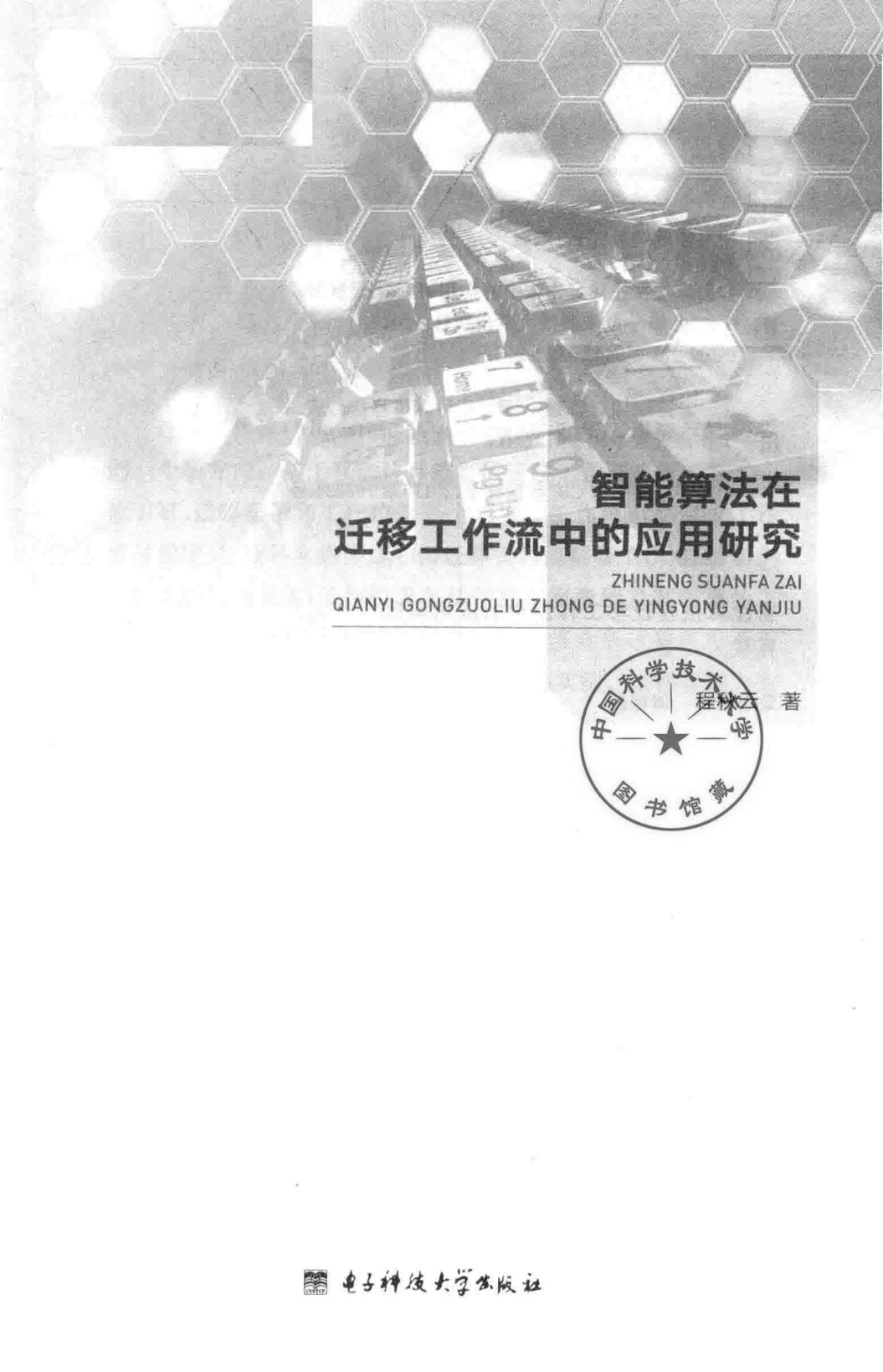
智能算法在 迁移工作流中的应用研究

ZHINENG SUANFA ZAI
QIANYI GONGZUOLIU ZHONG DE YINGYONG YANJIU

程秋云 著



电子科技大学出版社



智能算法在 迁移 workflow 中的应用研究

ZHINENG SUANFA ZAI

QIANYI GONGZUOLIU ZHONG DE YINGYONG YANJIU



著



电子科技大学出版社

图书在版编目(CIP)数据

智能算法在迁移 workflow 中的应用研究 / 程秋云
著. — 成都: 电子科技大学出版社, 2017.8
ISBN 978-7-5647-4851-7

I. ①智… II. ①程… III. ①计算机算法—应用—workflow 管理系统—研究 IV. ①F273-39

中国版本图书馆 CIP 数据核字(2017)第 182234 号

智能算法在迁移 workflow 中的应用研究

程秋云 著

策划编辑 谭炜麟

责任编辑 谭炜麟

出版发行 电子科技大学出版社

成都市一环路东一段 159 号电子信息产业大厦九楼 邮编 610051

主 页 www.uestcp.com.cn

服务电话 028-83203399

邮购电话 028-83201495

印 刷 河南承创印务有限公司

成品尺寸 145mm×210mm

印 张 7.5

字 数 193 千字

版 次 2017 年 11 月第一版

印 次 2017 年 11 月第一次印刷

书 号 ISBN 978-7-5647-4851-7

定 价 26.00 元

版权所有,侵权必究

前言

迁移工作流(Migrating Workflow)是近年工作流管理研究的一个新方向。基于移动计算范型的迁移工作流系统包括工作流引擎、迁移实例和工作位置三个要素。工作流引擎完成工作流过程定义、迁移实例生成和多迁移实例协调等工作;迁移实例是任务的执行主体;工作位置包括停靠站服务器及其管理的工作机网络,是迁移实例的运行场所。工作流业务过程根据业务目标的复杂程度被映射为一个或多个迁移实例。迁移实例为完成任务将根据自己的旅行图在工作位置间迁移。

停靠站服务器作为迁移实例的运行支撑平台和工作机网络的负载分配器,必须有良好的服务主体优选机制,以提高工作机网络的性能和资源利用率。

本书根据迁移工作流的概念模型和迁移工作流系统参考框架,基于遗传算法、Pareto 优化方法和自适应免疫算法,对停靠站服务器的服务主体优选机制进行了研究,提出了使工作位置多资源负载均衡的算法。

首先,本研究体系建立了服务主体优选的数学模型,构造了问题的多目标函数,并利用遗传算法对迁移域中的工作机状态进行编码;其次,遗传算法的优化结果可能引起正在运行的迁移实例脱离其迁移域,本书中提出了编码固定点的概念,以保证相应迁移实例的服务主体不发生变化;最后,给出了适应迁移工作

流特点的初始种群的构造方法以及相应的选择、交叉、变异等操作算子。

但是,上述方法把多目标问题单一化,得到的解可能引起较多的迁移实例进行二次迁移,反而增加系统负载。书中继而采用 Pareto 方法对算法进行了优化:首先,使用 Pareto 基于排序的适应度函数代替原来的适应度函数,并使用与此适应度函数相匹配的选择算子;其次,在算法过程中增加了 Pareto 最优解集和精英保留策略,并给出了排除非法解和选择决策方案的算法过程。同时,为避免遗传算法易于陷入局部最优和收敛速度较慢的问题,利用自适应免疫算法进行优化,给出了自适应交叉、变异操作、提取疫苗、局部最优判断、疫苗接种和免疫选择的具体方法。通过实验对算法的性能进行了分析和验证。

本书在撰写过程中得到了我的导师——山东大学韩芳溪副教授的指导,郑州航空工业管理学院计算机学院李玲玲教授、陈建辉副教授对本书的编写提出了很多宝贵建议。电子科技大学出版社对本书的出版给予了大力支持。在此表示衷心的感谢!本书涉及较多的专业知识和技术领域,在撰写过程中参考和引用了大量国内外相关领域研究成果和资料,在此向相关作者致以诚挚的谢意。由于我们的水平有限,书中难免有不当之处,敬请读者批评指正。

程秋云

2017 年 9 月

目 录

第 1 章 绪 论

1.1 研究背景和意义	001
1.2 研究内容与现状	002
1.3 创新点	004
1.4 本书的组织结构	005

第 2 章 智能算法概述

2.1 遗传算法	007
2.1.1 生物进化概述	007
2.1.2 遗传算法基本概念	008
2.1.3 遗传算法的应用	013
2.1.4 多目标优化问题	017
2.2 生物免疫系统	022
2.2.1 生物免疫学发展史	022
2.2.2 免疫系统与免疫学的关系	025
2.2.3 基于生物免疫系统的概念	026
2.2.4 生物免疫系统分类	027
2.2.5 生物免疫系统的组成	029
2.2.6 生物免疫系统的特点	036
2.2.7 免疫系统的机制	040
2.3 人工免疫系统	044
2.3.1 人工免疫系统模型	044

2.3.2 免疫算法	045
2.3.3 人工免疫系统的应用领域	051
2.4 免疫危险理论	055
2.4.1 自我-非我模式	055
2.4.2 感染的一非我模式	056
2.4.3 危险模式	057
2.4.4 危险理论在 SNS 中的表示	059
2.4.5 危险理论与传统理论的应用比较	061
2.4.6 危险理论研究现状	062
2.5 本章小结	062
第3章 移动 Agent 技术	
3.1 Agent 概述	063
3.1.1 什么是 Agent	063
3.1.2 Agent 特性	064
3.1.3 Agent 的类型	065
3.1.4 Agent 与对象	066
3.2 移动 Agent 技术	066
3.2.1 移动 Agent 概念	067
3.2.2 移动 Agent 的特征	068
3.2.3 移动 Agent 的主要优点	070
3.2.4 移动 Agent 的发展趋势	073
3.3 移动 Agent 的模型	074
3.3.1 移动 Agent 的理论模型	074
3.3.2 移动 Agent 的生命周期	075
3.3.3 移动 Agent 的交互	076
3.3.4 移动 Agent 通信语言	081

3.4 移动 Agent 的迁移	082
3.5 移动 Agent 系统	086
3.5.1 移动 Agent 系统的结构	086
3.5.2 典型的移动 Agent 系统	088
3.5.3 移动 Agent 系统中的关键技术	092
3.5.4 移动 Agent 系统的安全	094
3.5.5 移动 Agent 的技术标准	099
3.6 移动 Agent 的命名和创建	101
3.6.1 依赖位置的命名服务	101
3.6.2 独立于位置的命名服务	101
3.6.3 Agent 的创建	102
3.7 移动 Agent 的应用领域	102
3.8 本章小结	105
第 4 章 workflow 管理系统	
4.1 workflow 概述	106
4.1.1 workflow 的起源	106
4.1.2 workflow 定义	111
4.2 workflow 管理系统的概念	113
4.2.1 workflow 管理研究的发展历史	113
4.2.2 workflow 管理系统的定义	117
4.2.3 workflow 管理系统的分类	119
4.2.4 workflow 管理系统的优势	122
4.3 workflow 管理系统参考模型	124
4.3.1 workflow 管理系统体系结构	124
4.3.2 WfMC 的 workflow 参考模型	127



4.4 workflow 技术现状与发展趋势	130
4.5 本章小结	141
第5章 迁移 workflow 管理系统	
5.1 迁移 workflow 的产生	142
5.2 迁移 workflow 的模型和框架	144
5.2.1 迁移 workflow 模型	145
5.2.2 迁移 workflow 系统框架	148
5.3 迁移 workflow 的体系结构	150
5.3.1 迁移 workflow 引擎	150
5.3.2 停靠站服务器	152
5.3.3 工作位置	154
5.3.4 迁移实例	155
5.4 迁移 workflow 域的组织	158
5.4.1 工作流域及其组织模式	158
5.4.2 工作流域的静态组织	158
5.5 迁移 workflow 的执行机制	160
5.5.1 迁移 workflow 的执行过程	160
5.5.2 停靠站服务器的任务派遣机制	162
5.5.3 迁移实例的迁移规划	163
5.6 迁移 workflow 系统中的目录服务	167
5.6.1 迁移 workflow 对目录服务的要求	167
5.6.2 迁移 workflow 系统中的目录服务模型及其体系 结构	168
5.7 迁移 workflow 的安全现状	172
5.7.1 workflow 系统安全	172

5.7.2 迁移工作流管理系统的安全	174
5.8 一个迁移工作流系统的例子	175
5.9 本章小结	177
第6章 基于遗传算法和 Pareto 优化的服务主体优选机制	
6.1 负载均衡介绍	178
6.1.1 负载均衡的概念	178
6.1.2 网络负载均衡常用的方法	179
6.2 问题描述	182
6.3 服务主体优选的数学模型	182
6.4 遗传算法设计	183
6.4.1 迁移实例的编码	183
6.4.2 染色体种群初始化	184
6.4.3 适应度函数设计	185
6.4.4 选择操作	185
6.4.5 交叉操作	185
6.4.6 变异操作	186
6.5 Pareto 优化	187
6.5.1 适应度函数优化	187
6.5.2 选择操作优化	188
6.5.3 Pareto 解集和精英保留策略	188
6.5.4 排除非法解	189
6.5.5 选择决策方案	190
6.6 Pareto 遗传算法过程	190
6.7 算法举例及性能分析	191
6.7.1 算法举例	191

6.7.2 算法性能分析	193
6.8 本章小结	195
第7章 基于自适应免疫算法的服务主体优选	
7.1 自适应免疫遗传算法	196
7.1.1 识别抗原	197
7.1.2 初始化和选择操作	197
7.1.3 自适应交叉操作	197
7.1.4 自适应变异操作	198
7.1.5 提取疫苗	198
7.1.6 局部最优判断	198
7.1.7 疫苗接种和免疫选择	199
7.2 自适应免疫遗传算法过程	199
7.3 算法举例及性能分析	201
7.3.1 算法举例	201
7.3.2 算法性能分析	202
7.4 本章小结	203
第8章 总结与展望	204
参考文献	206

第 1 章 绪 论

1.1 研究背景和意义

workflow management is a computer supported collaborative work one of the important fields. [1] workflow management is in recent years in the computer application field of development most rapid of several new technologies, has been widely applied to office automation, CAD, file management, email, directory management, groupware application, BPR (Business Process Reengineering) and structural system definition tools etc., this makes workflow management system more and more to be valued by different application field users. workflow management system in many CSCW application software becomes more and more popular, demand continuously increases, at the same time workflow management thought has been widely applied to enterprise management and computer application of various aspects. but in application field expansion at the same time, also to workflow management put forward new requirements. [2]

migrating workflow (Migrating Workflow) is in recent years workflow management research of a new direction. [3] migrating workflow very well solved traditional workflow lack of flexibility, self-adaptability, normativity etc., obviously improved traditional workflow management model of flexibility and distribution, especially suitable for need to transfer large data and need to transfer large remote service distributed concurrent workflow process. [4]

in the mobile computing paradigm migrating workflow system, gave a migrating workflow management system framework, the framework by a migrating workflow management machine and several building friendly trust relationship local area network组成. migrating workflow management machine execute workflow engine, each local area network all contain a stop station server and several with its connected workstation network. migrating instance MI (Migration Instance) migrating



到某工作节点的目的是执行所携带的任务。^[5]停靠站服务器是 MI 的工作位置,它接受 MI 的迁移查询和迁移请求,并且在 MI 到达后为 MI 提供 workflow 服务和资源服务。停靠站服务器的核心工作在于如何为迁移实例分配工作机。此时应考虑以下三个方面。^[6]

①任务分解机制。工作机由于能力的限制,只能提供部分资源和服务,因此迁移实例迁移到停靠站服务器后,需要由停靠站服务引擎把 MI 在该节点执行的任务分解为多个子任务。

②服务主体(工作机)优选机制。当有多个工作机都能提供迁移实例 MI 所需的资源和服务时,服务引擎要提供优选算法,为迁移实例选择更合适的工作机。

③任务派遣机制。当迁移实例需要在多个工作机上执行任务时,服务引擎要根据任务特性而提供高效的任务派遣机制,特别要考虑多个子任务之间的串行与并行特性。

其中,服务主体的优选机制的好坏将直接影响系统资源的利用率及系统的性能。

目前,迁移 workflow 管理系统仍然处在实验室的原型研究阶段,其要成为下一代 workflow 管理系统还需要各个方面具体功能的研究和实现。针对停靠站服务器的具体设计及功能的实现的研究引起了越来越多的关注。

1.2 研究内容与现状

本书的主要目标是完善停靠站服务器的服务主体优选机制,使停靠站服务器能为迁移实例 MI 在它的迁移域中为其找到一台最适合的工作机执行任务,以维持迁移域的负载均衡,提



高系统资源的利用率。

对于抽取权因子的模糊综合评判算法^[7],由于其权因子确定的主观性和模糊算子的多样性,使该算法在实际使用中难以操作,并且此算法不能保证系统负载均衡,从而导致整个系统的资源利用率不高。

由于在迁移工作流中,迁移实例的服务主体必须首先满足它的资源和服务的需求,才能作为迁移实例的迁移目标,所以网络中常用的几种负载均衡算法在此是不适用的。

对于轮转调度算法^{[8][9]},因为每个迁移实例的迁移域都不相同,所以轮转根本就无法进行。而最小连接度算法^{[8][10]}由于没有考虑到任务大小,使得系统无法辨别迁移实例引起的负载(不同迁移实例引起的负载差别很大),也不适用。目标地址散列调度和源地址散列调度只要求所映射的服务器是可用且未超载就可以发送,反而会引起负载分布不均。且几种算法属于简单静态算法,既不能准确反映工作机负载的动态变化,也不能区分不同迁移实例占用的 CPU 资源和 I/O 资源的差异,无法实现系统多种资源的负载平衡。

本书首先从迁移工作流的特点入手,采用了遗传算法和 Pareto 优化的思想,提出了基于 Pareto 遗传算法的服务主体优选机制。同时根据迁移实例对资源的不同需求实现了多资源的负载均衡,并避免了调度倾泻,使迁移域中资源得到充分的应用,系统达到负载均衡。

其次,采用自适应免疫算法对遗传算法进行优化,保证种群的多样性,防止算法陷入局部最优,避免了未成熟收敛和收敛速度较慢的问题,提高遗传算法整体的优化性能。最后,模拟了实

验,证明了算法的有效性。

1.3 创新点

本书在基于移动计算范型的迁移工作流研究的基础上,重点对停靠站服务器的服务主体优选机制进行了研究,提出了一个基于 Pareto 优化的遗传算法优选机制;接着,在该算法的基础上,结合人工免疫算法,提出了自适应免疫算法的优选机制。该算法可以使系统负载均衡,提高系统资源利用率。

本书主要创新点如下。

①在迁移工作流中引入遗传算法。利用遗传算法搜索全局最优解的特点,确保算法的结果最优,并能有效避免调度倾泻。根据迁移实例必须在其迁移域中工作的特点,提出了固定点的概念,然后给出了适合迁移工作流特点的初始种群构造方法和选择、交叉、变异操作的算法过程。

②用 Pareto 方法对多目标函数进行优化。很好地适应了迁移工作流中不同迁移实例的迁移域不尽相同的特点,使决策者能够依据当前迁移实例运行状态选择最优解。

③提出了排除非法解和选择决策方案的算法过程,完成了从 Pareto 最优解集中系统自动选择决策的过程。

④采用自适应免疫算法对遗传算法进行优化,给出了自适应交叉、变异操作、提取疫苗、局部最优判断、疫苗接种和免疫选择的具体方法,避免了遗传算法易于陷入局部最优和收敛速度较慢的问题。



1.4 本书的组织结构

第1章论述了研究背景和意义,阐述了迁移工作流的研究现状,最后给出了本书的创新点。

第2章先介绍了遗传算法的起源、基本概念、应用以及多目标优化方法,接着简要介绍了生物免疫系统的概念、组成、特点和免疫机制,目的在于使大家对生物免疫系统有个全面的认识。在此基础上阐述了人工免疫系统模型、免疫算法和应用领域,并对免疫危险理论的表示、研究现状进行了概述。

第3章介绍了 Agent 和移动 Agent 的基本概念、特征和优点,阐述了移动 Agent 的模型和移动性,详细分析了移动 Agent 系统的结构、关键技术、安全和技术标准,给出了移动 Agent 的命名和创建方法。最后,根据移动 Agent 的特点和属性,分析了移动 Agent 的应用领域,为后续章节的研究工作做铺垫。

第4章介绍了工作流的起源、定义, workflow 系统的发展历史、定义、分类和优势,详细描述了 workflow 管理系统的体系结构和 WfMC 的工作流参考模型,分析了 workflow 的技术现状和发展趋势。

第5章介绍了迁移工作流的产生,描述了迁移工作流模型、系统框架和工作流域的组织模式,阐述了迁移工作流的组成部分——迁移工作流引擎、停靠站服务器、工作位置、迁移实例等的体系结构,详细介绍了迁移工作流的执行过程,重点分析了停靠站服务器任务派遣机制。最后,对目录服务和安全现状进行了概述,并给出了一个迁移工作流系统的例子。

第6章首先介绍负载均衡的概念和网络负载均衡的常用方法。其次建立了服务主体优选问题的数学模型,得到多目标函

数,并对多目标函数进行遗传算法设计,使用与迁移 workflow 相适应的编码、交叉、变异等操作算子。接着利用 Pareto 方法对遗传算法适应度函数进行优化,引入 Pareto 解集的保持机制和精英保留策略,给出了排除非法解和选择决策方案的算法过程,给出了 Pareto 遗传算法的具体算法过程和流程图。最后,对 Pareto 遗传算法进行了实验测试,证明了算法的有效性;从终止条件的选择、遗传算法优点的发挥、多资源负载均衡的实现和调度倾泻的避免等方面对算法的性能进行了详细的分析。

第 7 章基于人工免疫理论,提出自适应免疫遗传算法,详细介绍了自适应交叉、变异操作,给出了提取疫苗、局部最优判断、疫苗接种和免疫选择的具体方法,绘制了自适应免疫遗传算法流程图。最后,对自适应免疫算法进行了实验测试,和 Pareto 遗传算法进行了比较分析,有效避免了遗传算法在求解优化问题时所出现的易于陷入局部最优和出现早熟收敛现象,证明了算法的有效性;从免疫记忆、自我调节、抗体多样性等方面对算法的性能进行了详细的分析。

第 8 章是总结与展望。本书将智能算法用于服务主体优选策略中,并用 Pareto 方法对多目标问题进行优化,充分利用了智能算法搜索全局最优解、保持个体的多样性的特点,保证迁移实例的工作位置最优。本书中算法的编码是单倍体的,也可以仿照高等生物采用双倍体遗传。双倍体遗传提供了一种记忆以前有用的基因块的功能,使得当前环境再次变为以前发生过的情况时,物种会很快适应。