

- 中文核心期刊
- 中国科技核心期刊
- RCCSE中国核心学术期刊

ISSN 1671-6833
CN 41-1339/T
CODEN ZDXGAN

郑州大学学报

工 学 版

JOURNAL OF ZHENGZHOU UNIVERSITY
ENGINEERING SCIENCE

2017.6

第38卷 第6期 (卷终)

Vo1.38 No.6

2017年11月出版

ISSN 1671-6833



目 次

信息工程

- 基于可能性条件偏好网络的交互式遗传算法及其应用····· 孙晓燕, 朱利霞, 陈 杨(1)
- 无线 Ad Hoc 网络中双向传输容量的研究 ····· 陈昌海, 李娜娜, 王 静, 李振锋(6)
- 可重构网络中的一种新型端到端服务供应模型····· 马 丁, 庄 雷, 兰巨龙, 慕小武(11)
- 单云服务提供者环境下的随机化属性保护研究····· 李拴保(17)
- 基于喷泉码及 CDMA 扩频的数字水印盲算法 ····· 张卫党, 吴 琳(23)
- 基于混合差分演化的网络入侵检测算法····· 王耀光, 陈伟权, 吴镇邦, 秦 勇, 黄 翰(29)
- 精英反向学习的单纯形交叉布谷鸟搜索算法····· 林要华, 王李进(33)

机械与电气工程

- 考虑公差原则的三维公差分析方法研究····· 赵凤霞, 金少搏, 李纪峰(39)
- 计及线路损耗的峰谷分时电价优化模型····· 赵国生, 詹天乐, 李 博(45)
- 基于 Möbius 变换的复有理圆弧样条 ····· 任两品, 薛均晓, 张朝阳, 王定标(50)
- 基于 FPGA 的直流电机伺服控制系统设计与实现····· 常怡萍, 刘 琴, 杨 蕾(54)
- 考虑顾客时间紧迫度的生鲜电商配送路径优化问题 ·····
····· 崔 岩, 张子祥, 时 新, 王晓亮, 王振锋(59)
- 基于 Mean Shift 聚类的多级阈值化方法 ····· 程 全, 刘晓青, 刘玉春, 王志良(64)
- 石油钻井工程预警技术发展概述····· 李 广, 张 衡, 王 杰, 朱晓东, 岳彩通(70)
- 具备直流故障隔离能力的新型 MMC 拓扑研究 ····· 张清枝, 李志辉(74)

化工与材料工程

- 路面免碾压贫混凝土基层工作性研究····· 盛燕萍, 孙仕伟, 赵澍林, 周辉丽, 陈拴发(79)
- 改性凹凸棒石对水溶液中 Pb(II) 吸附性能研究 ····· 张 婕, 蔡 丽, 王景涛, 陈卫航(83)
- 典型 MoSi₂ 发热元件物相与显微结构分析 ·····
····· 黄文江, 周 颖, 李 斯, 马成良, 史幸福, 王世界(88)
- LiAlCl₄/γ-Al₂O₃ 催化歧化制备二甲基二氯硅烷的研究 ····· 徐文媛, 姚成姣, 徐永昌, 洪三国(92)

CONTENTS

【Informational Engineering】

- Probabilistic Conditional Preference Network Assisted Interactive Genetic Algorithm and Its Application
..... **SUN Xiaoyan, ZHU Lixia, CHEN Yang (1)**
- Bidirectional Transmission Capacity in Wireless Ad Hoc Networks
..... **CHEN Changhai, LI Nana, WANG Jing, LI Zhenfeng (6)**
- A Novel End-to-end Service Provisioning Model in Reconfigurable Network
..... **MA Ding, ZHUANG Lei, LAN Julong, MU Xiaowu (11)**
- Research on Randomization Attribute Protection in Single Cloud Service Provider **LI Shuanbao (17)**
- Blind Digital Watermarking Algorithm Based on Fountain Code and CDMA Spread Spectrum
..... **ZHANG Weidang, WU Lin (23)**
- Network Intrusion Detection Algorithm Based on Hybrid Differential Evolution Algorithm
..... **WANG Yaoguang, CHEN Weiquan, WU Zhenbang, QIN Yong, HUANG Han (29)**
- Elite Opposition-based Learning Based Simplex Crossover Cuckoo Search Algorithm **LIN Yaohua, WANG Lijin (33)**

【Mechanical Engineering and Electrical Engineering】

- Research on the Method for 3D Tolerance Analysis Considering Tolerance Principle
..... **ZHAO Fengxia, JIN Shaobo, LI Jifeng (39)**
- Time-of-use Price Optimization Model Considering Line Loss **ZHAO Guosheng, ZHAN Tianle, LI Bo (45)**
- Circular Arc Representation Based on Möbius Transformation
..... **REN Liangpin, XUE Junxiao, ZHANG Chaoyang, WANG Dingbiao (50)**
- Design and Implementation of Motor Servo Control System Based on FPGA
..... **CHANG Yiping, LIU Qin, YANG Lei (54)**
- Fresh Agricultural Product E-commerce Distribution Routing Problem Considering Time Demand of Customer
..... **CUI Yan, ZHANG Zixiang, SHI Xin, WANG Xiaoliang, WANG Zhenfeng (59)**
- Based on the Mean Shift Clustering Multilevel Threshold Method
..... **CHENG Quan, LIU Xiaoqing, LIU Yuchun, WANG Zhiliang (64)**
- A Review: Pre-warning System of Oil-Drilling Engineering
..... **LI Guang, ZHANG Heng, WANG Jie, ZHU Xiaodong, YUE Caitong (70)**
- Research on the Topology of MMC with DC Fault Isolating Capability **ZHANG Qingzhi, LI Zhihui (74)**

【Chemical Engineering and Material Engineering】

- Research on Workability of Compaction-free Lean Concrete Base
..... **SHENG Yanping, SUN Shiwei, ZHAO Shulin, ZHOU Huili, CHEN Shuanfa (79)**
- Study on Absorption of Pb(II) from Aqueous Solution by Modified ATP
..... **ZHANG Jie, CAI Li, WANG Jingtao, CHEN Weihang (83)**
- Study on Phase Composition and Microstructure of Typical MoSi₂ Heating Elements
..... **HUANG Wenjiang, ZHOU Ying, LI Si, MA Chengliang, SHI Xingfu, WANG Shijie (88)**
- The Theoretical Research of Dimethyldichlorosilane Catalyzed by LiAlCl₄/γ-Al₂O₃
..... **XU Wenyuan, YAO Chengjiao, XU Yongchang, HONG Sanguo (92)**

文章编号: 1671-6833(2017)06-0001-05

基于可能性条件偏好网络的交互式遗传算法及其应用

孙晓燕, 朱利霞, 陈 杨

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221008)

摘 要: 根据用户实施的人机交互行为而隐式地获取用户偏好的交互式进化优化算法, 可有效减轻用户疲劳, 提高个性化搜索或推荐的效率。但是, 已有研究没有考虑用户交互行为和偏好的不确定性, 影响了对用户偏好的拟合精度以及基于该偏好表达的进化搜索。针对该问题, 提出基于可能性条件偏好网络的交互式遗传算法, 以刻画用户交互行为和偏好的不确定性, 并提高算法的搜索性能。首先, 采用交互时间表示交互行为, 考虑交互行为的不确定性, 给出交互时间可信度的定义, 并基于该定义给出了用户不确定偏好的表达函数; 其次, 利用可信交互时间和偏好函数, 定义了用户对评价对象的偏好权重, 并利用该权重, 设计(更新)可以定量表示用户不确定偏好的可能性条件偏好网络, 以更好地拟合用户偏好; 然后, 结合评价不确定性和可能性条件偏好网络, 提出了改进的个体适应值估计策略, 以更好地引导搜索; 最后, 将所提算法应用于图书个性化搜索中, 结果表明了算法搜索的可靠性和高效性。

关键词: 交互式遗传算法; 不确定性; 可能性条件偏好网络; 个性化搜索

中图分类号: TP181

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.001

0 引言

在当前信息社会中, 个性化搜索、个性化服务等已成为网络发展、数据管理和挖掘的首要任务。实质上, 个性化信息获取是优化问题。但是, 目前从优化角度解决该类问题的研究成果还较少。在个性化信息获取中, 人机交互贯穿整个过程, 如用户兴趣建模和跟踪等^[1]。若将用户兴趣模型与进化优化算法结合, 则可望提高用户搜索效率, 而这恰好符合交互式进化优化算法应用的范畴^[2]。交互式进化优化算法将人的智能与进化优化搜索过程进行结合, 可有效解决个性化产品设计、电力系统优化、房屋装修、网页布局、音乐创作及图像检索等问题^[3-8]。

孙晓燕等^[9]借鉴个性化搜索中用户兴趣建模的研究成果, 考虑用户搜索的评价对象属性间的相互关联关系, 提出了基于交互行为和条件偏好网络的交互式进化优化算法, 并用于解决个性化搜索问题。然而, 在该算法框架中, 没有考虑用户认知或偏好评价的不确定性。笔者在文献[9]的基础上, 研究含认知不确定性的隐式评价交互式进化优化算法。

由于鼠标滚动信息也可反映用户的偏好^[10], 因此, 可利用浏览时间和鼠标滚动次数定义交互时间和评价偏好的不确定性函数。因条件偏好网络(conditional preference networks, CP-nets)无法反映节点以及节点之间存在的偏好不确定性, 文献[11]提出了一种新的偏好网络——可能性网络(probabilistic networks), 它能灵活地表达不确定偏好。在此基础上, Cornelio^[12]和 Bigot 等^[13]考虑到动态不确定和噪音特性, 进一步构建动态概率 CP-nets, 即每个依赖关系都以一定的概率存在, 从而灵活刻画偏好关系。笔者在此基础上利用 PCP-nets (probabilistic conditional preference networks) 可灵活表示不确定偏好。

1 基于交互行为的用户偏好不确定性表示

在电子商务搜索系统中, 用户的交互行为能很好地体现用户的兴趣偏好。若仅仅利用交互时间的长短去刻画用户对某搜索对象的偏好度难以反映用户交互行为中的所有信息。如用户对第 i 个商品 X_i 和第 j 个商品 X_j 的交互浏览时间关系为 $t(X_i) > t(X_j)$, 并不能说明用户对 X_i 的兴趣度一

收稿日期: 2017-05-18; 修订日期: 2017-08-04

基金项目: 国家自然科学基金资助项目(61473298)

作者简介: 孙晓燕(1978—), 女, 江苏徐州人, 中国矿业大学教授, 博士, 主要从事智能优化研究, E-mail: xysun78@126.com.

定比对 X_j 的兴趣度大,有可能是 X_i 的页面较长,增加了用户的浏览时间,那么 $t(X_i)$ 和 $t(X_j)$ 就有一定的可信度. 所以,笔者结合鼠标滚动次数,考虑具有可信度的交互行为.

1.1 交互行为的可信度

考虑 4 类交互行为: ① 点击浏览行为 A_1 ; ② 保存行为 A_2 ; ③ 收藏/标签行为 A_3 ; ④ 决策/购买行为 A_4 . 偏好关系为 $A_1 < A_2 < A_3 < A_4$.

采用交互时间刻画上述 4 种交互行为,并利用鼠标滚动次数定义交互时间的可信度. 对搜索对象 X_i , 设用户交互时间为 $t_j(X_i)$, 鼠标滚动次数为 $n_j(X_i)$, $j = 1, 2, 3, 4$; $i = 1, 2, \dots, L_j$, 其中 L_j 表示对象集合包含的元素个数. 融合可信度的交互行为量化表达如式(1)所示:

$$t'_j(X_i) = t_j(X_i) e^{-\frac{n_j(X_i)}{\beta}}, \quad (1)$$

式中, β 为控制条件可信程度变化的参数.

1.2 含交互不确定性的偏好表示

基于上述交互行为可信度表示,进一步给出用户对搜索对象 X_i 的偏好表示. 用户对于特别喜欢和特别不喜欢的搜索对象往往给出的评价越可信; 而对于不关心的对象, 评价的可信性则相对较小. 那么用户偏好不确定度定义如下所示:

$$g_j(X_i) = e^{-\frac{[t'_j(X_i) - t_0]^2}{\gamma}}, \quad (2)$$

式中, γ 为可变参数; t_0 为平均浏览时间.

$$t_0 = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^4 t'_j(X_i). \quad (3)$$

当 $t'_j(X_i) < t_0$ 时,说明用户对商品不感兴趣; 当 $t'_j(X_i) \geq t_0$ 时,说明用户对商品感兴趣. 所以,给出用户对 X_i 的偏好函数如式(4)所示:

$$f_j(X_i) = \begin{cases} ke^{-g_j(X_i)}, & t'_j(X_i) \geq t_0; \\ ke^{-1} g_j(X_i), & t'_j(X_i) < t_0, \end{cases} \quad (4)$$

式中, k 为调节参数.

用户对搜索对象 X_i 的偏好比重 w_{ji} , $i \in \{1, 2, \dots, L_j\}$, $j \in \{1, 2, 3, 4\}$, 如式(5)所示:

$$w_{ji} = t'_j(X_i) f_j(X_i). \quad (5)$$

对式(5)进行归一化处理:

$$w_{ji} = \frac{w_{ji}'}{\sum_{j=1}^4 \sum_{i=1}^{L_j} w_{ji}'}. \quad (6)$$

2 基于不确定偏好权重的可能性条件偏好网络的构建

2.1 可能性条件偏好网络的初始化构建

采用文献[9]的方法构建初始化条件偏好有

向图,不再赘述,重点是可能性条件偏好表的构造. 设 PCP-nets 的偏好有向图共包含多层节点,其中第 l 层有 m_l 个节点变量,记为 $x_1^l, x_2^l, \dots, x_{m_l}^l$, 第 k 个变量 x_k^l 具有 n_k 个取值 $x_{k,1}^l, x_{k,2}^l, \dots, x_{k,n_k}^l$, 那么 $x_{k,j}^l$ 在所有父节点可能取值组合条件下的偏好可能性设共有 M_{kj}^l 种组合,各偏好可能性取值记为 $\alpha_{k,j,1}^l, \alpha_{k,j,2}^l, \dots, \alpha_{k,j,M_{kj}^l}^l$. 对于父节点具有明确偏好属性取值的 $\alpha_{k,j,p}^l = 1$, 其余的 $\alpha_{k,j,p}^l$ 在 $[0, 1)$ 随机取值.

2.2 可能性条件偏好网络的更新

首先,可能性条件偏好表的更新. 通过式(4)可获得决策变量的偏好度. 设决策变量的第 m 个取值 $x_{l,m}$ 在 A_j 交互行为下出现次数为 n_j^{lm} , 那么 $x_{l,m}$ 的偏好度如式(7)所示:

$$f(x_{l,m}) = \frac{\sum_{j=1}^4 f_j(x_l) n_j^{lm}}{\sum_{j=1}^4 n_j^{lm}}. \quad (7)$$

其次,条件偏好有向图的更新. 由式(6)可得用户在交互行为 A_j 下对评价对象 X_i 的偏好权重为 w_{ji} . 设当前搜索群体中,决策变量 x_l 在 A_j 交互行为下的取值共有 n_j^l 个,决策变量的偏好权重如式(8)所示:

$$w(x_l) = \frac{\sum_{j=1}^4 w_{ji} n_j^l}{\sum_{j=1}^4 n_j^l}. \quad (8)$$

根据式(8)计算偏好权重,更改支配关系.

3 基于可能性条件偏好网络的交互式遗传算法

3.1 基于可能性条件偏好网络的个体适应值的估计

首先,考虑条件偏好有向图中决策变量的层级分布贡献度. x_j 支配的决策变量越多, $class(x_j)$ 越小. 根据文献[9]可知,决策变量 x_j 的贡献度为 $C_1(x_j) = 2^{(class(x_j)-1)} class(x_j)$, 若某决策变量未出现,则设该贡献度值为 $C_1(x_j) = 1$. 结合偏好可能度,借鉴贝叶斯推理的链式公式,对于进化个体 $X_i = \{x_{i1}, x_{i2}, \dots, x_{in_i}\}$ 的适应值估计策略如式(9):

$$F(X_i) = \prod_{j=1}^n 2^{(class(x_j)-1)} class(x_j) \alpha(x_{j,i_j}^{class(x_j)}). \quad (9)$$

3.2 算法步骤

Step1: 根据实际问题进行编码,并设定遗传操作的各种参数;

Step2: 基于用户输入关键词信息, 历史搜索信息和社会群体搜索结果, 获得初始化种群;

Step3: 根据 2.1 节方法, 构建初始化可能性条件偏好网络;

Step4: 根据 3.1 节方法, 估计当前进化种群的个体适应值, 选择 N 个适应值较大的进化个体提交给用户;

Step5: 用户实施交互行为, 若找到满意解, 则算法终止, 并输出满意解, 否则, 转 Step6;

Step6: 记录浏览时间和鼠标滑动次数, 根据第 1 节计算用户对评价对象的偏好权重, 并根据 2.2 节方法更新可能性条件偏好网络;

Step7: 对当前种群实施选择, 交叉和变异操作, 生成新种群, 转 Step4.

4 算法的应用

4.1 实验背景

将本文算法应用到文献 [9] 所构建的图书个性化搜索系统中, 如图 1 所示. 通过算法的比较, 验证本文算法的有效性. 为了说明本文算法 IGA-PCP (probabilistic conditional preference network assisted interactive genetic algorithm) 的整体性能, 将其与文献 [9] 所提算法 IGA-CP (conditional preference network assisted interactive genetic algorithm) 以及传统的交互式遗传算法 IGA (interactive genetic algorithm) 进行比较.



图 1 图书搜索系统界面

Fig. 1 Interface of book search system

4.2 参数设置

采用二进制形式编码, 编码串按照决策属性分块. 以当当网心理学图书为例, 将心理学图书属性分为 7 类, 每一类属性下有若干取值, 本实验中共 46 个, 需要 19 位二进制码. 实验中, $\beta = 6$, $k = 5$, $\gamma = 4$, 3 种比较算法均采用轮盘赌选择, 单点交叉和变异, 且交叉和变异概率分别为 0.6、0.1, 种群规模设为 8.

4.3 实验结果与分析

4.3.1 偏好不确定函数表示的合理性

鼠标滚动次数以及可信时间之间的关系如表 1 所示. 可信时间与所定义的偏好度函数关系如图 2 所示.

从表 1 可知, 鼠标滚动次数与用户实际交互时间的相互关系与本文所提可信时间相符, 表明笔者所提时间可信度的合理性. 由图 2 可知, 式 (4) 所定义的偏好度函数随着可信时间的增加而非线性增加, 符合实际情况, 表明了偏好度函数定义的合理性.

表 1 IGA-PCP 用户浏览行为

Tab. 1 IGA-PCP user browsing behavior

图书个体	用户实际交互时间/s	鼠标滚动次数	可信时间/s
儿童发展与教育心理学	15	2	11
自闭症	18	4	9
学前儿童心理学	20	6	7
特殊儿童及青少年心理学	50	8	13
魔鬼心理学	4	5	2
重口味心理学	16	4	8
女人气场心理学	1	1	1
女性心理成长自疗课	20	9	4

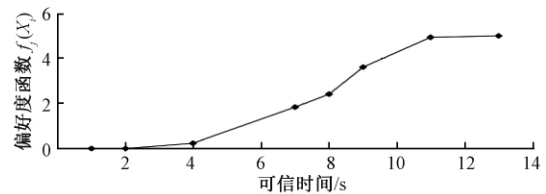


图 2 偏好度函数 $f_j(X_i)$ 与可信时间的关系

Fig. 2 Relation between preference function $f_j(X_i)$ and credible time

4.3.2 对比实验

3 种比较算法, 实验结果如图 3 所示. 实验针对 84 类图书共 4 420 本, 在所有的图书中找到“特殊儿童的问题行为干预”, 记录实验数据如图 3 所示: ①IGA-PCP 平均互异个体的比率最高, 这说明了本文算法 IGA-PCP 能有效提高搜索的多样性. ②本文算法的搜索时间比传统 IGA 大约少一半, 并且明显小于 IGA-CP 算法搜索时间. ③IGA-PCP 用户实施交互操作行为的次数平均为 4.0, IGA-CP 平均为 7.6, 传统的 IGA 平均为 7.7, 由此可以看出, 本文算法 IGA-PCP 可有效减轻用户疲劳.

为了说明所提算法的整体性能, 现对励志类图书进行搜索, 实验针对 24 类图书共 1 159 本, 实验结果如表 2 所示. 从表 2 可以看出, 本文算法不仅有效减轻用户疲劳, 而且可有效提高算法多

样性. 总体来说,本文算法明显优于其他两种算法.

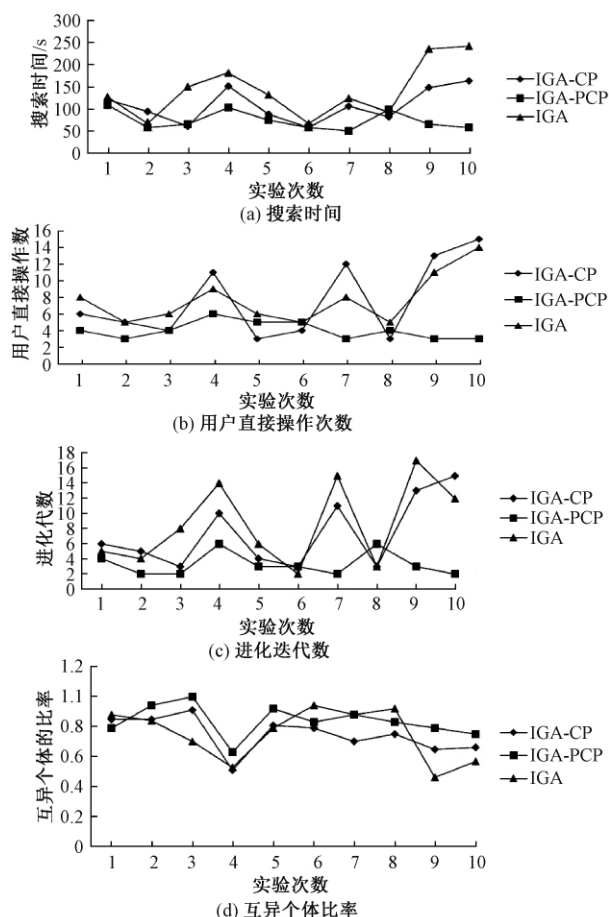


图3 84类心理学图书搜索实验结果

Fig.3 The results of the experiment of eighty four kinds of psychology books

表2 24类励志图书搜索实验结果

Tab.2 The results of the experiment of twenty-four kinds of inspirational books

算法	平均进化代数	平均搜索时间/s	用户交互次数	互异个体比率
IGA-PCP	5.3	93.4	5.9	0.848
IGA-CP	7.8	112.6	7.8	0.689
IGA	9.6	126.0	7.2	0.669

5 结束语

用户在搜索过程中,用户的交互行为能够有效地反映用户偏好信息,但由于用户认知的不确定性,反映到交互行为上也具有不确定性. 针对此情况,笔者考虑了交互行为的不确定性刻画,以及基于该不确定性刻画的用户不确定偏好的描述,在用户偏好建模方面,基于所考虑的不确定性偏好,提出了一种可能性条件偏好网络拟合用户评价偏好的方法,给出了适应值估计策略,并将算

法应用到图书搜索系统中,实验结果表明了该算法的可行性与有效性.

如何利用其他用户的信息构建多用户偏好模型,并进行有效集成和更新,以实现网络环境下的信息动态更新并提高搜索效率,将是下一步要研究的问题.

参考文献:

- [1] QIAN X M, FENG H, ZHAO G S, et al. Personalized recommendation combining user interest and social circle [J]. IEEE transactions on knowledge data engineering, 2013, 26(7): 1763-1777.
- [2] 孙晓燕. 高级交互式遗传算法理论与应用 [M]. 北京: 科学出版社, 2010: 7-18.
- [3] SORM D, RIMCHAROEN S. Web page template design using interactive genetic algorithm [C] // International Computer Science and Engineering Conference (ICSEC). Nakorn Pathom, Thailand: IEEE, 2013: 201-206.
- [4] 杨胡萍, 李威仁, 左士伟, 等. 基于改进遗传算法的电力系统无功优化 [J]. 郑州大学学报(工学版), 2015, 36(6): 66-69.
- [5] LIU X. Color mapping design from image to 3D product model [J]. Journal of mechanical engineering, 2009, 45(10): 222-227.
- [6] GARCIA-HERNANDEZ L, ARAUZO-AZOFRA A, SALAS-MORERA L, et al. Recycling plants layout design by means of an interactive genetic algorithm [J]. Intelligent automation and soft computing, 2013, 19(3): 457-468.
- [7] KOGA S, INOUE T, FUKUMOTO M. A proposal for intervention by user in interactive genetic algorithm for creation of music melody [C] // International Conference on Biometrics and Kansei Engineering. Tokyo, Japan: IEEE, 2013: 129-132.
- [8] DASS M V, ALI M M, ALI M R. Image retrieval using interactive genetic algorithm [C] // International Conference on Computational Science and Computational Intelligence. Las Vegas, USA: IEEE, 2014: 215-220.
- [9] 孙晓燕, 陆宜娜, 巩敦卫, 等. 基于 CP-nets 的偏好感知交互式遗传算法及其个性化搜索 [J]. 控制与决策, 2015, 30(7): 1153-1161.
- [10] CLAYPOOL M, LE P, WASED M, et al. Implicit interest indicators [C] // Proceedings of the 6th International Conference on Intelligent User Interfaces. New York, USA: IEEE, 2001: 33-40.
- [11] BENAMOR N, DUBOIS D, GOUIDER H, et al. Possibilistic networks: a new setting for modeling prefer-

- ences [M]. Berlin, Germany: Springer, 2014: 1 – 7.
- [12] CORNELIO C, GOLDSMITH J, MATTEI N, et al. Dynamic probabilistic CP-nets [C] // Multidisciplinary Workshop on Advances in Preference Handling. Chicago, USA: IEEE, 2013: 1 – 7.
- [13] BIGOT D, FAGIER H, MENGIN J, et al. Probabilistic conditional preference networks [J]. Computer science, 2013, 19: 72 – 81.

Probabilistic Conditional Preference Network Assisted Interactive Genetic Algorithm and Its Application

SUN Xiaoyan, ZHU Lixia, CHEN Yang

(School of Information and Control Engineering, China University of Mining & Technology, Xuzhou 221008, China)

Abstract: Interactive evolutionary algorithms with user preference implicitly extracted from interactions of user were more powerful in alleviating user fatigue and improving the exploration in personalized search or recommendation. However, the uncertainties in user interactions and preferences have not been considered in the previous research, which might greatly impact the reliability of the extracted preference model, as well as the effective exploration of the evolution with that model. Therefore, an interactive genetic algorithm with probabilistic conditional preference networks (PCP-nets) was proposed, in which, the uncertainties were further figured out according to the interactions, and a PCP-net was designed to depict user preference model with higher accuracy by involving those uncertainties. First, the interaction time was adopted to mathematically describe the relationship between the interactions and user preference, and the reliability of the interaction time was further defined to reflect the interactive uncertainty. The preference function with evaluation uncertainty was established with the reliability of interaction time. Second, the preference weights on each interacted object were assigned on the basis of preference function and reliability. With these weights, the PCP-nets were designed and updated by involving the uncertainties into the preference model to improve the approximation. Third, a more accurate fitness function was delivered to assign fitness for the individuals. Last, the proposed algorithm was applied to a personalized book search and its superiority in exploration and feasibility was experimentally demonstrated.

Key words: interactive genetic algorithm; uncertainty; possibilistic conditional preference networks; personalized search

无线 Ad Hoc 网络中双向传输容量的研究

陈昌海^{1,3}, 李娜娜², 王 静³, 李振锋³

(1. 电子科技大学 光互联网与移动信息网络研究中心, 四川 成都 611731; 2. 郑州轻工业学院 计算机与通信工程学院, 河南 郑州 450002; 3. 四川工程职业技术学院, 四川 德阳 618000)

摘 要: 目前对于基于双向传输信道的无线 Ad Hoc 网络传输容量的研究还很少, 主要原因是无法得到传输容量的闭合表达式, 无法准确地研究各种网络参数对传输容量的影响. 针对这一情况, 使用随机几何理论建立无线 Ad Hoc 网络双向传输模型, 并对瑞利衰落信道时双向传输成功概率和双向传输容量的闭合表达式进行了理论推导. 数值仿真结果表明: 单次传输的成功概率受多个网络参数的影响, 实际网络中通信密度的增加会导致通信成功概率的迅速下降; 双向传输容量主要受到网络允许的最大通信失败概率的影响, 当最大通信失败概率较小时, 双向传输容量随着最大通信失败概率的增加而显著增大, 但是此时通信质量也会下降, 因此需要选择合适的最大通信失败概率以便兼顾通信质量和双向传输容量.

关键词: Ad Hoc; 双向传输; 传输容量; 最大通信失败概率; 通信对

中图分类号: TN911.22

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.002

0 引言

无线 Ad Hoc 网络的传输容量^[1] (TC, transmission capacity) 被定义为具有最大通信失败概率约束的最大成功通信密度, 也就是说, TC 实质上是在满足一定通信质量时网络能够在单位面积上实现的最大成功通信数量. 多年来, 人们试图从多个方面^[2] 对无线 Ad Hoc 网络的 TC 进行研究, 目前广泛为研究者们所认可的研究方法是使用随机几何理论^[3] 对网络中发送机 (TX, transmitter) 和接收机 (RX, receiver) 在网络中的位置分布进行建模. 使用这一方法, 研究者们从不同方面对 TC 进行了研究, 比如干扰消除^[4]、多天线^[5]、保护区域^[6] 等, 并取得了较为显著的研究成果.

前期的研究主要都是基于网络中的通信是单向传输^[7-8] 这一条件, 认为只要任意 TX 发送的信号到达对应的 RX 处的信干比 (SIR, signal-to-interference ratio) 大于信干比门限值就能够实现信号的成功传输^[9]. 但是实际上的通信系统往往是双向通信, 就算是单向通信的通信系统也需要反馈信号来进行信道初始化、信道反馈、接收确认

和路由信息传递等. 文献 [10] 首次提出了在无线 Ad Hoc 网络 TC 的研究中使用双向信道, 并对双向信道对 TC 的影响进行了研究, 但是只给出了 TC 的上下界的范围, 没有得到 TC 的闭合表达式, 无法给出 TC 与网络参数之间的具体关系, 因此无法准确的描述各种网络参数变化对 TC 的影响. 针对这一问题, 笔者研究瑞利衰落信道情况下的无线 Ad Hoc 网络的双向传输容量 (BTC, bidirectional transmission capacity), 使用随机几何理论对无线 Ad Hoc 网络进行建模, 并推导 BTC 的闭合表达式. 在此网络中, 对于任意 TX 和 RX 构成的通信对来说, 一次成功的双向通信包括从 TX 到 RX 和 RX 到 TX 两次通信, 只有当这两次通信都成功时才能认为这次双向通信成功.

1 系统模型

使用随机几何理论对无线 Ad Hoc 网络进行建模, 网络处于二维平面上且没有规模限制, 在网络中存在两组通信终端, 一组为 TX, 另一组为 RX, 任意 TX 都有唯一的一个 RX 与之相对应, 并构成一个通信对. 根据随机几何理论, 网络中所有

收稿日期: 2017-02-10; 修订日期: 2017-06-20

基金项目: 国家自然科学基金资助项目 (61271257)

作者简介: 陈昌海 (1980—), 男, 吉林东丰人, 电子科技大学讲师, 博士, 主要从事新进无线通信技术及信号处理等方面研究, E-mail: chenchanghai163@163.com.

TX_s(或 RX_s) 在网络中的位置服从密度为 λ 泊松点过程 (PPP, poisson point process)。

使用时隙 Aloha 随机接入协议^[11],在任意时隙内,任意通信对以概率 p 随机接入信道进行通信。根据随机几何理论中 PPP 具有平稳性^[12],在任意时隙内,所有进行通信的 TX(或 RX) 在网络中的位置服从密度为 $p\lambda$ 的 PPP。网络中所有的通信对使用频分双工进行双向通信,即任意 TX 使用相同的频率 f_t 发送数据给对应的 RX,并且任意 RX 使用相同的频率 f_r 发送数据给对应的 TX。在任意时隙内,随机选定一组通信对并用编号 0 来表示,网络中的其他通信对用编号 i ($i = 1, 2, 3, \dots$) 来表示,对编号为 0 的通信对的通信性能统计特性的研究结果将与其它通信对相同。

在某一时刻内,任意 TX₀ 与 RX₀ 之间的距离为 d ; TX _{i} 与 RX₀ 之间的距离用 $X_{t,i}$ 表示;信道功率衰落系数用 $H_{t,i}$ 表示; RX _{i} 与 TX₀ 之间的距离用 $X_{r,i}$ 表示;信道功率衰落系数用 $H_{r,i}$ 表示。由于无线 Ad Hoc 网络是干扰受限系统^[13],笔者不考虑随机噪声的影响,定义信道为瑞利衰落,信号功率衰减因子服从参数为 τ 的指数分布,路径损耗因子为 α ($\alpha > 2$),则 RX₀ 处接收到的信号为

$$S_r = PH_{t,0}d^{-\alpha} + \sum_{i=1,2,3,\dots} PH_{t,i}X_{t,i}^{-\alpha}. \quad (1)$$

在 TX₀ 处接收到的信号为

$$S_t = PH_{r,0}d^{-\alpha} + \sum_{i=1,2,3,\dots} PH_{r,i}X_{r,i}^{-\alpha}, \quad (2)$$

式中: P 为任意 TX(或 RX) 的发送信号功率。因此可以得到 RX₀ 处的 SIR_r 为

$$SIR_r = \frac{H_{t,0}d^{-\alpha}}{\sum_{i=1,2,3,\dots} H_{t,i}X_{t,i}^{-\alpha}}. \quad (3)$$

TX₀ 处的 SIR_t 为

$$SIR_t = \frac{H_{r,0}d^{-\alpha}}{\sum_{i=1,2,3,\dots} H_{r,i}X_{r,i}^{-\alpha}}. \quad (4)$$

定义 SIR 的最低门限值为 β ,在任意 TX 或 RX 处的 SIR 一旦低于 β 就会导致通信失败,因此 RX₀ 接收信号成功概率可以表示为

$$S_r = Pr(SIR_r \geq \beta). \quad (5)$$

TX₀ 处接收信号成功概率可以表示为

$$S_t = Pr(SIR_t \geq \beta). \quad (6)$$

只有当 RX₀ 和 TX₀ 处都成功接收信号时,编号为 0 的通信对 (TX₀ 和 RX₀) 才成功地进行了一次双向通信,因此双向通信失败概率为

$$Q = 1 - S_r S_t. \quad (7)$$

为了保证网络的通信质量,一般要求双向通信

失败概率要控制在一个较小的值^[14],定义 ε 为网络允许的最大通信失败概率,则通过对 $Q = \varepsilon$ 进行求解可以得到对应的最大通信对密度 λ^ε 。参考单向传输网络 TC 的定义,笔者定义 BTC 为

$$C_B = (1 - \varepsilon) p \lambda^\varepsilon, 0 < \varepsilon < 1. \quad (8)$$

2 双向传输的统计特性研究

根据 RX₀ 处成功通信概率的定义(5),把式(3)代入公式(5)并进行整理可以得到:

$$S_r = Pr(H_{t,0} \geq \beta d^\alpha I_t), \quad (9)$$

式中,

$$I_t = \sum_{i=1,2,3,\dots} H_{t,i}X_{t,i}^{-\alpha}. \quad (10)$$

定义符号 $E(\cdot)$ 为对随机参数求数学期望,由于 $H_{t,0}$ 服从参数为 τ 的指数分布,则公式(9)可以变为

$$S_r = E\left(\int_{\beta d^\alpha I_t}^{\infty} \tau e^{-\tau h} dh\right) = L_{I_t}(\tau \beta d^\alpha), \quad (11)$$

式中: $L_{I_t}(s)$ 表示随机变量 I_t 的拉普拉斯变换。

定理 1: 随机变量 $I_t = \sum_{i=1,2,3,\dots} H_{t,i}X_{t,i}^{-\alpha}$ 的拉普拉斯变换表达式为

$$L_{I_t}(s) = e^{-\lambda p \pi s^\gamma E(H_{t,i}^\gamma) \Gamma(1-\gamma)}, \quad (12)$$

式中: $\gamma = 2/\alpha$; $\Gamma(x)$ 是伽玛函数。

证明: 为了能够推导随机变量 I_t 的拉普拉斯变换 $L_{I_t}(s)$, 首先把网络区域限制在以 RX₀ 为圆心并且半径为 r 的圆形区域,此时如果只有 k 个 TX 存在于此区域,则 I_t 的拉普拉斯变换可以表示为

$$L_{I_t,k}(s) = E(e^{-s \sum_{i=1}^k H_{t,i}X_{t,i}^{-\alpha}}) = [E(e^{-s H_{t,i}X_{t,i}^{-\alpha}})]^k. \quad (13)$$

由于 TX_s 服从密度为 λp 的 PPP,则在面积为 πr^2 的区域内存在 k 个 TX_s 的概率为

$$P(k, \text{TX}_s) = \frac{(\lambda p \pi r^2)^k}{k!} e^{-\lambda p \pi r^2}. \quad (14)$$

把式(13)与(14)结合在一起,可以得到任意数量 TX 位于面积为 πr^2 的圆形区域内时的 I_t 的拉普拉斯变换的期望值为

$$\begin{aligned} L_{I_t}(s) &= \sum_{k=0}^{\infty} \frac{(\lambda p \pi r^2)^k}{k!} e^{-\lambda p \pi r^2} [E(e^{-s H_{t,i}X_{t,i}^{-\alpha}})]^k = \\ &= e^{\lambda p \pi r^2 [E(e^{-s H_{t,i}X_{t,i}^{-\alpha}}) - 1]} \cdot \sum_{k=0}^{\infty} \frac{(\lambda p \pi r^2 E(e^{-s H_{t,i}X_{t,i}^{-\alpha}}))^k}{k!} e^{-\lambda p \pi r^2 E(e^{-s H_{t,i}X_{t,i}^{-\alpha}})} = \\ &= e^{\lambda p \pi r^2 [E(e^{-s H_{t,i}X_{t,i}^{-\alpha}}) - 1]}. \end{aligned} \quad (15)$$

由于任意 TX_i 随机分布在此圆形区域内, $X_{t,i}$ 的概率密度函数为

$$f_X(x) = 2x/r^2, 0 \leq x \leq r. \quad (16)$$

把公式(16)代入式(15)可以得到:

$$L_{I_i}(s) = e^{\lambda p \pi [E(\int_0^{\tau} e^{-s H_{t,i} x} - \alpha 2 x dx) - 1]}. \quad (17)$$

当 $r \rightarrow \infty$ 时,公式(17)变为

$$L_{I_i}(s) = e^{-\lambda p \pi s \gamma E(H_{t,i}^\gamma) \Gamma(1-\gamma)}. \quad (18)$$

定理 1 证明完毕.

定理 2: 在双向对称传输信道系统模型中,无线 Ad Hoc 网络的双向传输容量为

$$C_B = \frac{-(1-\varepsilon) \ln(1-\varepsilon)}{2\gamma\beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (19)$$

证明: 由于 $H_{t,i}$ 服从参数为 τ 的指数分布,因此可以得到:

$$E(H_{t,i}^\gamma) = \int_0^\infty h^\gamma \tau e^{-\tau h} dh = \tau^{-\gamma} \Gamma(1+\gamma). \quad (20)$$

把式(20)代入式(12)可以得到随机变量

$I_i = \sum_{i=1,2,3,\dots} H_{t,i} X_{t,i}^{-\alpha}$ 的拉普拉斯变换表达式为

$$L_{I_i}(s) = e^{-\lambda p \pi s \gamma \tau^{-\gamma} \Gamma(1+\gamma) \Gamma(1-\gamma)}. \quad (21)$$

由于

$$\Gamma(1+\gamma) \Gamma(1-\gamma) = \gamma \pi \csc(\gamma\pi). \quad (22)$$

把式(22)代入式(21),则随机变量 $I_i =$

$\sum_{i=1,2,3,\dots} H_{t,i} X_{t,i}^{-\alpha}$ 的拉普拉斯变换表达式变为

$$L_{I_i}(s) = e^{-\lambda p \gamma \pi^2 \tau^{-\gamma} \csc(\gamma\pi) s^\gamma}. \quad (23)$$

式(11)是单向传输成功概率的表达式,把式(23)代入式(11)可以得到单向传输成功接收信号的

$$S_r = L_{I_i}(\tau \beta d^\alpha) = e^{-\lambda p \gamma \beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (24)$$

由于网络中的 TX_s 和 RX_s 具有相同的分布,信道传输参数也具有相同的统计特性,因此使用与推导 RX_0 处成功接收信号概率相同的推导过程可以得到:

$$S_t = S_r = e^{-\lambda p \gamma \beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (25)$$

把式(24)和式(25)代入双向传输失败概率定义式(7),可以得到双向通信失败概率为

$$Q = 1 - e^{-\lambda p \gamma \beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (26)$$

令 $Q = \varepsilon$ 并对式(26)进行求解,可以得到对应于最大通信失败概率 ε 时的最大通信对密度 λ^ε 为

$$\lambda^\varepsilon = \frac{-\ln(1-\varepsilon)}{2p\gamma\beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (27)$$

最后,把最大通信对密度 λ^ε 式(27)代入 BTC 的定义式(8),可以得到无线 Ad Hoc 网络的双向传输容量为

$$C_B = \frac{-(1-\varepsilon) \ln(1-\varepsilon)}{2\gamma\beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (28)$$

定理 2 证明完毕.

3 分析和讨论

这一章将对前面推导的统计特性结果进行数值仿真分析,如果没有特别的说明,仿真参数如表 1 所示.

表 1 仿真参数

Tab.1 Simulation parameters

名称	符号	数值
随机接入概率	p	0.001
路径衰减因子	α	4
SIR 的最低门限值	β	10
TX_i 与 RX_i 之间的距离/m	d	20

由于 $S_t = S_r$, 可以认为单向传输成功通信概率与 RX_0 处成功接收信号概率相同. 参考 RX_0 处成功接收信号概率的式(20)可知,成功通信概率主要受通信对密度 λ 、信干比最低门限值 β 、路径衰减因子 α 、随机接入概率 p 和通信对的传输距离 d 等参数影响. 如图 1 所示,单向传输的成功通信概率随着通信对密度 λ 的增大而迅速减少,这是因为通信对密度 λ 的增大使得同一时隙内通信的通信对增多,由于使用相同的通信频率,通信对之间的干扰会显著增多,使得通信失败的概率迅速增大. 另外可以看到,单向传输的成功通信概率随着信干比最低门限值的减少而增大,但是此时的通信质量也会随之降低.

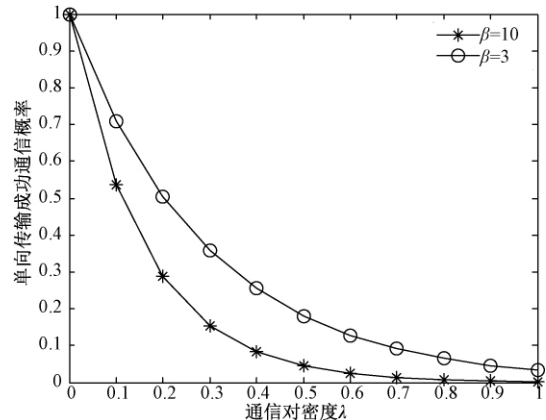


图 1 单向传输成功通信概率

Fig.1 Probability of successful single-way transmission

对于单向传输系统,使用笔者定义的网络模型和参数,可以得到单向传输系统的 TC 为

$$C_s = \frac{-(1-\varepsilon) \ln(1-\varepsilon)}{\gamma\beta^\gamma d^2 \pi^2 \csc(\gamma\pi)}. \quad (29)$$

如图 2 所示,单向传输系统 TC 和双向传输系统 BTC 首先随着最大通信失败概率 ε 的增大而增大,后随着最大通信失败概率 ε 的增大而减少,并且存在一个最优 ε 使得网络的 TC 和 BTC

达到最大值. 参考 BTC 的定义公式(8)可以知道, 网络的 BTC 主要由最大通信对密度 λ^ε 和成功通信概率 $1 - \varepsilon$ 的乘积决定. 当最大通信失败概率 ε 较小时, ε 的增大会使得 λ^ε 有较为明显的增大, 而此时成功通信概率 $1 - \varepsilon$ 的减小相对较少, 最终使得网络的 BTC 增大. 这说明, 适当的降低网络的通信质量对于提高网络的 BTC 有较为明显的作用. 当最大通信失败概率 ε 较大时, ε 的增大对 λ^ε 变化的影响开始变得不明显, 而此时成功通信概率 $1 - \varepsilon$ 的减小变得相对较大, 最终使得网络的 BTC 减少. 这种情况说明降低网络的通信质量并不一定能够提高网络的 BTC, 有时候会取得相反的效果.

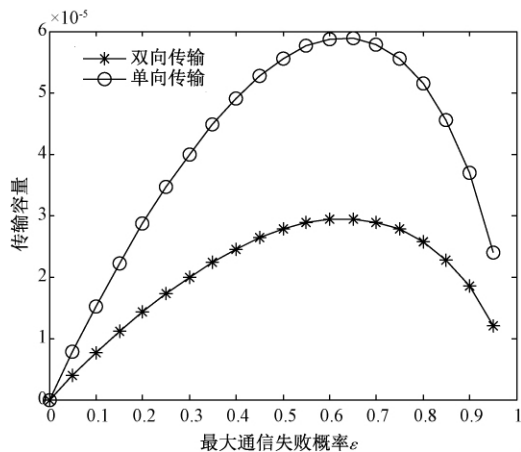


图2 传输容量与最大通信失败概率的关系图

Fig. 2 Transmission capacity versus the maximum outage probability

从上面的分析可知, 可以通过控制网络中的最大通信失败概率 ε 来实现对 BTC 的控制. 一般网络中, ε 往往是一个较小的值, 此时增大 ε 会使得网络的 BTC 增加, 但是会导致通信质量的下降, 因此如何在保证通信质量和提高网络 BTC 之间权衡^[11] 是实际网络参数设置时首先要考虑的.

另外从图2中可以看到, 当最大通信失败概率 ε 保持不变时, 最大双向传输的 BTC 要比单向传输的 TC 小, 但是两者随 ε 变化的趋势是一致的. 这一结果说明, 前期论文对网络参数与单向传输 TC 关系的研究结果对双向传输的 BTC 同样有用.

4 结论

对于无线 Ad Hoc 网络来说, 双向传输系统模型比单向传输系统模型更加贴近实际网络, 因此对于 BTC 的研究结果更加具有实际应用价值. 笔者的研究表明, 无线 Ad Hoc 网络的 BTC 主

要受通信对密度和最大通信失败概率影响, 适当的提高最大通信失败概率可以使得网络的 BTC 有较大的提高, 代价则是通信质量会下降; 但是当最大通信失败概率较大时, 反而会使得网络的 BTC 减少, 说明为了兼顾通信质量和 BTC 需要合理的设置最大通信失败概率. 本文的讨论仅仅局限于双向对称信道, 而实际网络中有很多是非对称双向信道, 此时双向传输的带宽和通信速率都会不同, 网络参数对 BTC 的影响还有待于进一步研究.

参考文献:

- [1] WEBER S, ANDREWS J G, JINDAL N. The effect of fading, channel inversion, and threshold scheduling on ad hoc networks [J]. IEEE transactions on information theory, 2007, 53(11): 4127–4149.
- [2] LEE J, JEFFREY G A, HONG D. Spectrum-sharing transmission capacity with interference cancellation [J]. IEEE transactions on communications, 2014, 61(1): 76–86.
- [3] GAMAL A E, KIM Y H. Network information theory [M]. Cambridge: Cambridge University, 2012.
- [4] HUANG S, RAMAMOORTHY A. On the multiple-unicast capacity of 3-source, 3-terminal directed acyclic networks [J]. IEEE/ACM transactions on networking, 2015, 22(1): 285–299.
- [5] HUNTER A M, ANDREWS J G, WEBER S, et al. Capacity scaling of ad hoc networks with spatial diversity [J]. IEEE transactions on wireless communications, 2008, 7(12): 58–71.
- [6] HASAN A, ANDREWS J G. The guard zone in wireless ad hoc networks [J]. IEEE transactions on wireless communications, 2007, 6(1): 897–906.
- [7] KOYLUOGLU O, KOKSAL C E, GAMAL H E. On secrecy capacity scaling in wireless networks [J]. IEEE transactions on information theory, 2016, 58(1): 3000–3015.
- [8] WEBER S, YANG X, ANDREWS J, et al. Transmission capacity of wireless ad hoc networks with outage constraints [J]. IEEE transactions on information theory, 2011, 51(12): 4091–4102.
- [9] 陈昌海, 钟伟, 李振锋, 等. 重叠无线 Ad Hoc 网络中进度密度的研究 [J]. 科技通报, 2015, 31(5): 229–233.
- [10] VAZE R, TRUONG K, HEATH R, et al. Two-way transmission capacity of wireless ad-hoc networks [J]. IEEE transactions on wireless communications, 2010, 10(6): 1966–1975.

- [11] KIM D, PARK S, JU H, et al. Transmission capacity of full-duplex-based two-way ad hoc networks with ARQ protocol [J]. IEEE transactions on vehicular technology, 2015, 63(7): 3167–3183.
- [12] ALI O, CARDINAL C, GAGNON F. Performance of optimum combining in a poisson field of interferers and Rayleigh fading channels [J]. IEEE transactions on wireless communications, 2010, 9(8): 2461–2467.
- [13] 娄思佳, 何俊, 姬瑞龙. 对 Ad Hoc 网络通信干扰的可行性分析 [J]. 火力与指挥控制, 2015, 40(6): 103–106.
- [14] 李娜娜, 陈昌海, 甘勇, 等. 分布式网络中进度密度的研究 [J]. 郑州大学学报(工学版), 2014, 35(3): 51–54.

Bidirectional Transmission Capacity in Wireless Ad Hoc Networks

CHEN Changhai^{1,3}, LI Nana², WANG Jing³, LI Zhenfeng³

(1. Research Center of Optical Internet and Mobile Information Networks, University of Electronic Science and Technology of China, Chengdu 611731, China; 2. School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China; 3. Sichuan Engineering Technical College, Deyang 618000, China)

Abstract: The research of the transmission capacity in wireless Ad Hoc networks based on bidirectional channels grew slowly. The closed-form expression could not be obtained. As a result, the relationship between the transmission capacity and the network parameters could not be expressed exactly. To solve the problem, the system model of bidirectional transmission in wireless Ad Hoc networks, was proposed to redefine the bidirectional transmission capacity based on the symmetrical bidirectional transmission channel, and to derive the probability of successful transmission and transmission capacity for bidirectional transmission. Numerical results showed that, the probability of successful transmission for single transmission was affected by some network parameters, and was mainly determined by the density of transmission pair with which the probability of successful transmission would increase. The bidirectional transmission capacity was mainly determined by the maximum outage probability. When the maximum outage probability was a small value, the bidirectional transmission capacity would increase with maximum outage probability, while the quality of communication would reduce, so how to choose a suitable maximum outage probability was very important in real networks.

Key words: Ad Hoc; bidirectional transmission; transmission capacity; maximum outage probability; density of transmission pair

可重构网络中的一种新型端到端服务供应模型

马 丁^{1,2}, 庄 雷¹, 兰巨龙³, 慕小武⁴

(1. 郑州大学 信息工程学院, 河南 郑州 450001; 2. 河南工业大学 信息科学与工程学院, 河南 郑州 450001;
3. 国家数字交换系统工程技术研究中心, 河南 郑州 450002; 4. 郑州大学 数学与统计学院, 河南 郑州 450001)

摘 要: 现有服务承载网的构建方法均未考虑用户请求所属的业务类别, 导致业务承载与服务供应缺少有效的串联. 针对该问题, 提出了一种新型的端到端服务供应模型, 从承载同类别业务的角度构建服务承载网, 通过在服务承载网上构建服务路径提供可定制的端到端服务. 为验证模型的可行性和有效性, 设计了一种基于跳数约束的服务承载网构建算法, 并使用经典的 Layered Graph 算法构建服务路径. 实验结果体现了在不同的负载压力下, 跳数约束对开销、收益、接受率、路径长度等性能评价指标的影响, 从而验证了模型和算法的有效性, 提供了性价比最优的服务承载网构建决策. 当跳数约束等于3时, 取得了与跳数约束等于4或5时相近的性能, 同时分别节省了约50%或75%的构建开销.

关键词: 可重构网络; 元能力; 服务承载网; 服务供应模型; 服务路径

中图分类号: TP393

文献标志码: A

doi: 10.3969/j.issn.1671-6833.2017.06.003

0 引言

随着互联网规模的增长以及应用范围的扩展, 端系统的多样化以及新兴通信模式的不断涌现, 互联网内的体系结构“僵化”^[1]、可扩展性差等缺点开始逐渐显现, 在互联网中增加新的功能和部署新的业务变得非常困难. 在当前互联网的体系结构中, 大多数复杂的网络功能和数据通信服务被放置在位于网络边缘的端系统之上, 而位于网络核心的路由器只是提供简单的存储转发服务^[2]. 这种端到端的数据传输模式虽然简化了互联网的设计, 但是却严重限制了其体系结构的灵活性, 对新的应用需求适应性极差.

针对上述问题, 可重构信息通信基础网络^[3]提供了相应的解决方案:

(1) 可重构网络在其体系结构中加入了对网络虚拟化技术的支持, 通过在共享的底层网络之上构建服务承载网 (service carrying network, SCN) 的方式为用户提供服务. 现阶段对服务承载网构建问题 (虚拟网映射问题) 的研究^[4-7], 主要考虑在资源有限的可重构网络中如何实现资源

的分配, 高效地将服务承载网拓扑和资源请求映射到底层网络上, 但是均未考虑用户对网络功能的需求.

(2) 可重构网络将网络功能分解为细粒度的网络功能单元, 称之为元能力 (atomic capability, AC). 利用具有可编程特性的路由器基础平台^[5], 元能力可以根据需要实现或部署在可重构核心网络的任意节点上. 对于用户的请求, 可重构管理服务器运行服务路由算法, 构建端到端的服务路径 (service path, SP), 使数据流依次经过所需的元能力实例, 从而提供可定制的服务. 目前针对可重构网络服务路径构建问题 (service path constructing problem, SPC) 的研究尚处在初始阶段, 在可编程网络、云计算、网络功能虚拟化等网络环境中, 现有的研究成果^[8-13]均是基于底层网络进行服务实例的部署与路径选择, 均未考虑建立中间的服务承载层.

基于上述分析, 笔者提出一种新型端到端服务供应模型, 其核心思想是基于元能力理论, 从承载同类型业务的角度对服务承载网构建问题重新建模, 将服务承载网构建为虚拟化的中间抽象层, 然

收稿日期: 2016-08-03; **修订日期:** 2017-02-21

基金项目: 国家重点基础研究发展计划 (“973”计划) (2012CB315901); 国家自然科学基金资助项目 (61379079); 河南省国际合作项目 (152102410021)

作者简介: 马丁 (1978—), 男, 河南郑州人, 河南工业大学讲师, 博士, 主要从事可重构网络技术研究, E-mail: martindingzz@163.com.

后提出了基于跳数约束的服务承载网构建算法,并结合 Layered Graph 算法对该模型进行仿真验证。

1 服务供应机制

在服务承载网构建问题中^[4-5],请求表示为网络拓扑的形式,链路是非有序的,节点之间只有资源需求的不同而没有顺序的差别,可重构管理服务器通过在底层网络上创建服务承载网提供端到端的服务。

当可重构网络的节点具有实现和部署网络功能的能力之后,请求可以表示为元能力序列和资源需求,体现了端到端服务多样化、可定制的特性。服务承载网由最终的服务供应形式转变为承载同类业务的虚拟化中间层,为服务路径构建提供一致性的网络拓扑和资源视图。服务供应的方式则转化为在服务承载网上构建服务路径来提供端到端的服务。这种新型服务供应模式的主要步骤如下:

- (1) 根据服务请求中的元能力需求,判断请求所属的业务类别;
- (2) 如果承载该类业务的服务承载网存在,则由其承载服务请求;如果不存在,则在底层网络之上创建一个满足该类业务需求的服务承载网,由其承载服务请求;
- (3) 在承载服务请求的服务承载网之上运行服务路径构建算法,最终建立一条从源节点到达目的节点的最优路径,使所需的元能力在该路径上被依次执行。

2 服务供应模型

2.1 可重构网络

可重构网络,即底层网络描述为带权无向图 $G^S = (N^S, L^S, A_N^S, A_L^S)$, 其中 N^S 和 L^S 分别表示底层网络的节点集合和链路集合; A_N^S 和 A_L^S 分别表示节点和链路所具有的资源属性集合。底层网络节点 $n^S \in N^S$ 拥有的资源属性包括: 计算能力 $C^S(n^S)$; 元能力实例集合 $AC^S(n^S) = \{ac_k^S \mid \text{节点 } n^S \text{ 能够提供元能力实例 } ac_k^S\}$; 元能力实例的处理时间 $d^S(ac_k^S(n^S))$; 使用单位节点资源需要的开销 $c^S(n^S)$ 。底层网络链路 $l^S \in L^S$ 拥有的资源属性包括: 带宽 $B^S(l^S)$; 通信时延 $d^S(l^S)$; 单位链路资源需要的开销 $c^S(l^S)$ 。如图 1 所示的可重构网络中,节点内的数字表示剩余计算能力; ac_1 等表示元能力实例; 链路上的数字表示剩余带宽; h 表示接入路径的跳数。

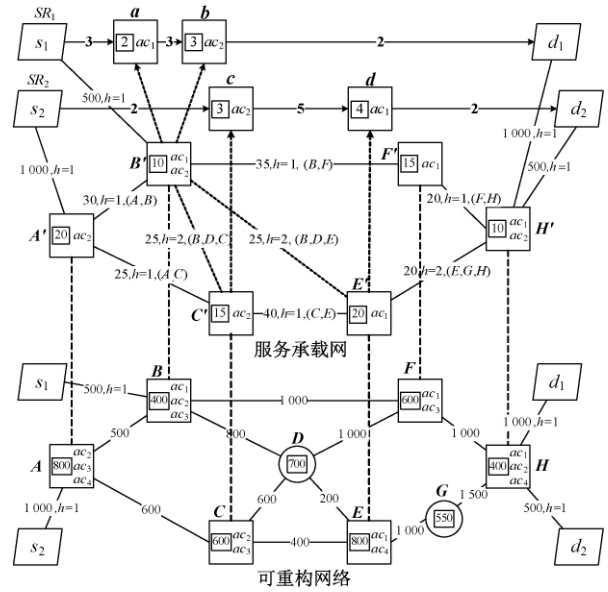


图 1 服务供应示例

Fig. 1 Example of service provisioning

2.2 服务请求

端系统的服务请求 (Service Request, SR) 描述为逻辑服务路径的形式,其拓扑标记为带权有向图 $G^R = (N^R, L^R, R_N^R, R_L^R)$, 其中 N^R 和 L^R 分别表示逻辑节点和逻辑链路的集合; R_N^R 和 R_L^R 分别表示逻辑节点和逻辑链路所需的资源属性集合。逻辑节点 $\forall n^R \in N^R$ 需求的资源属性包括: 元能力需求 $ac^R(n^R)$; 元能力的计算能力需求 $\mu(n^R)$; 使用单位节点资源需要支付的开销 $c^R(n^R)$ 。逻辑链路 $l^R \in L^R$ 需求的资源属性包括: 逻辑链路的带宽需求 $\mu(l^R)$; 使用单位链路资源需要支付的开销 $c^R(l^R)$ 。服务请求所需的元能力集合 $AC^R = \{ac_1^R, ac_2^R, \dots, ac_A^R\}$ 。如图 1 所示,服务请求 SR_1 、 SR_2 中,逻辑节点内部的 ac_1 和数字表示元能力和计算能力需求,逻辑链路上的数字表示带宽需求。

2.3 服务承载网构建问题

服务承载网描述为带权无向图 $G^V = (N^V, L^V, A_N^V, A_L^V)$, 其中 N^V 和 L^V 分别表示服务承载网的节点和链路集合; A_N^V 和 A_L^V 分别表示节点和链路所具有的资源属性集合。 AC^R 所属业务类别的候选节点集合表示为 $N_R^S = N^S - \{n^S \mid AC^S(n^S) \cap AC^R = \emptyset\}$ 。

服务承载网的构建过程分为两个步骤:

- (1) 从底层网络中选择候选节点集合 N_R^S , 建立 N^V 与 N_R^S 之间的映射 $M_N: N^V \rightarrow N_R^S, \forall n^V \in N^V, M_N(n^V) \in N_R^S$, 使服务承载网的每个节点映射在不同的底层节点上。同时建立 N^V 与 N_R^S 之间的逆映射 $M_N': N_R^S \rightarrow N^V, \forall n^S \in N_R^S, M_N'(n^S) \in N^V$, 使

服务承载网节点和底层网络节点形成一对一映射。然后为节点分配资源: 计算能力 $C^V(n^V) = \theta C^S(M_N(n^V))$, $0 < \theta \leq 1$; 元能力实例 $AC^V(n^V) = AC^S(M_N(n^V))$, 元能力的处理时间 $d^V(ac^V(n^V)) = d^S(ac^S(M_N(n^V)))$, 单位节点资源开销 $c^V(n^V) = \delta c^S(n^S)$, $0 < \delta \leq 1$ 。

(2) 在节点 $n_i^V, n_j^V \in N^V$ 之间构建链路, 建立 L^V 与 \mathbf{L}^S 之间的映射 $M_L: L^V \rightarrow \mathbf{L}^S, \forall (n_i^V, n_j^V) \in L^V, M_L(n_i^V, n_j^V) \subseteq \mathbf{L}^S(M_N(n_i^V), M_N(n_j^V))$ 。其中, \mathbf{L}^S 表示底层网络所有路径的集合; $P^S \in \mathbf{L}^S$ 表示一条底层网络的路径; $h(P^S)$ 表示路径 P^S 的跳数; $h(n_i^V, n_j^V)$ 表示链路 (n_i^V, n_j^V) 所映射路径的跳数; $B^S(P^S)$ 表示路径 P^S 的带宽, 为路径上链路带宽的最小值。然后, 为链路分配资源: $\forall n^V \in N^V, P^S \in M_L(l^V)$, 链路带宽 $B^V(l^V) = \varphi B^S(P^S)$, $0 < \varphi \leq 1$; 链路通信时延 $d^V(l^V) = \sum_{l^S \in P^S} d^S(l^S)$; 单位链路资源需要的开销 $c^V(l^V) = \psi h(P^S) c^S(l^S)$, $0 < \psi \leq 1$ 。

在图1中, 服务承载网链路 (B^V, E^V) 上的信息表示: 带宽为25, 所映射路径跳数为2, 路径为 (B, D, E) 。

2.4 服务路径构建问题

服务路径的构建过程分为两个步骤:

(1) 逻辑节点映射: $\mathbf{N}: N^R \rightarrow N^V, \forall n^R \in N^R, \mathbf{N}(n^R) \in N^V, ac^R(n^R) \in AC^V(\mathbf{N}(n^R))$ 。

(2) 逻辑链路映射: $\mathbf{L}: L^R \rightarrow \mathbf{L}^V, \forall (n_i^R, n_j^R) \in L^R, \mathbf{L}(n_i^R, n_j^R) \subseteq \mathbf{L}^V(\mathbf{N}(n_i^R), \mathbf{N}(n_j^R))$, $\mathbf{L}^V, P^V, h(P^V)$ 的表述与底层网络中的类似。

服务路径 SP 定义为

$$SP = \{s, (n_{i_1}^V, \mathbf{N}(n_{i_1}^R)), (n_{i_2}^V, \mathbf{N}(n_{i_2}^R)), \dots, (n_{i_\omega}^V, \mathbf{N}(n_{i_\omega}^R)), d\}.$$

ω 为服务路径中节点的数量, $\mathbf{N}(n_{i_k}^R)$ 是节点 $n_{i_k}^V$ 上逻辑节点到该节点映射的集合, 定义为 $\mathbf{N}(n_{i_k}^R) = \{(n^R \rightarrow n_{i_k}^V) \mid \mathbf{N}(n^R) = n_{i_k}^V\}$ 。

同时满足以下资源约束条件:

(1) $\forall n_{i_k}^V \in SP, \mathbf{N}(n_{i_k}^R) \neq \emptyset$,

$$C_{\text{avail}}^V(n_{i_k}^V) \geq \sum_{\forall n^R \rightarrow n_{i_k}^V \in \mathbf{N}(n_{i_k}^R)} \mu(n^R).$$

即服务路径的节点有足够的计算能力执行映射至其上的元能力实例, 其中 $C_{\text{avail}}^V(n_{i_k}^V)$ 表示节点 $n_{i_k}^V$ 的可用计算能力。

(2) $\forall l_{i_k}^V \in SP$,

$$B_{\text{avail}}^V(l_{i_k}^V) \geq \sum_{P^V \in \mathbf{L}(l_{i_k}^R), l_{i_k}^V \in P^V} \mu(l^R).$$

即服务路径链路的带宽满足逻辑链路的带宽

需求约束, 其中 $B_{\text{avail}}^V(l_{i_k}^V)$ 表示链路 $l_{i_k}^V$ 的可用带宽。

2.5 评价指标

2.5.1 服务承载网构建问题的评价指标

在底层网络上构建服务承载网的开销主要由分配计算资源和带宽资源产生的开销组成, 定义为

$$C(G^V) = \sum_{n^V \in N^V} c^V(n^V) C^V(n^V) + \sum_{l^V \in L^V} c^V(l^V) B^V(l^V). \quad (1)$$

2.5.2 服务路径构建问题评价指标

服务路径的长度定义为对应底层路径的跳数:

$$h(SP) = h(s, n_{i_1}^V) + \sum_{k=1}^{\omega-1} h(n_{i_k}^V, n_{i_{k+1}}^V) + h(n_{i_\omega}^V, d).$$

与文献[4]类似, 服务承载网在 t 时刻接受一个服务请求的收益定义为

$$R(G^R, t) = \sum_{n^R \in N^R} c^R(n^R) \mu(n^R) + \sum_{l^R \in L^R} c^R(l^R) \mu(l^R). \quad (2)$$

服务承载网在 t 时刻接受一个服务请求的开销定义为

$$C(G^R, t) = \sum_{n^R \in N^R} c^V(\mathbf{N}(n^R)) \mu(n^R) + \sum_{l^R \in L^R} \sum_{P^V \in \mathbf{L}(l^R)} \sum_{l^V \in P^V} c^V(l^V) \mu(l^R). \quad (3)$$

服务路径的总时延等于服务路径上元能力实例的处理时间与链路时延之和, 定义为

$$D(SP) = \sum_{k=1}^{\omega} \sum_{\forall n^R \rightarrow n_{i_k}^V \in \mathbf{N}(n_{i_k}^R)} d^V(ac^R(n_{i_k}^R)) + d^V(s, n_{i_1}^V) + \sum_{k=1}^{\omega-1} d^V(n_{i_k}^V, n_{i_{k+1}}^V) + d^V(n_{i_\omega}^V, d).$$

可重构网络构建服务承载网之后, 通过长期运营获取收益, 对服务承载网而言, 端到端服务供应的长期综合质量可以用以下指标度量。

(1) 服务承载网的长期平均运营收益:

$$R(G^R) = \lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T R(G^R, t)}{T}. \quad (4)$$

(2) 服务承载网的长期平均运营开销:

$$C(G^R) = \lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T C(G^R, t)}{T}. \quad (5)$$

(3) 平均服务请求接受率:

$$\lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T SR_{\text{success}}}{\sum_{t=0}^T SR}. \quad (6)$$

(4) 服务承载网的长期平均收益开销比定义为

$$\frac{R(G^R)}{C(G^R)} = \lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T R(G^R, t)}{\sum_{t=0}^T C(G^R, t)}. \quad (7)$$

(5) 平均服务路径长度:

$$\lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T h(SP)}{\sum_{t=0}^T SR_{\text{success}}}. \quad (8)$$

(6) 平均服务路径时延:

$$\lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T D(SP)}{\sum_{t=0}^T SR_{\text{success}}}. \quad (9)$$

式中: SR_{success} 表示成功接受的服务请求。

3 仿真与结果分析

3.1 实验环境设置

为了对所提出模型的可行性和性能进行评价,设计了基于跳数约束的服务承载网构建算法,在候选节点之间建立满足跳数约束 hops 的链路,并采用 Layered Graph^[8],简称 LG 算法构建服务路径,水平边权值设定为链路的 h 值,垂直边权值设定为 1。

与文献[7]和[13]类似,可重构网络和服务请求的设置如下:底层网络拓扑采用 GT-ITM 工具^[12]随机生成,包括 50 个节点和约 120 条链路,节点的计算能力和链路的带宽服从 [500, 1 000] 的均匀分布,每个节点拥有随机的 1 ~ 5 种不同类型的元能力。为简化问题,所有链路时延均设定为 10 个时间单位,所有元能力实例的处理时间均设定为 5 个时间单位,这样 LG 算法求出的最短路径也是最小时延路径。单位资源开销 $c^S(n^S)$ 、 $c^S(l^S)$ 、 $c^R(n^R)$ 、 $c^R(l^R)$ 均设定为 1.0;系数 θ 和 φ 设定为 0.1;系数 δ 和 ψ 设定为 1.0。

假设端系统服务请求的到达时间服从泊松分布,为测试算法在不同请求到达率下的表现,考虑平均 100 个时间单位内请求到达率的均值为 4、5、6、7、8 的 5 种情况。请求的业务类别设定为元能力 ac_1 、 ac_2 、 ac_3 的任意组合,元能力的计算能力需求服从 [0, 25] 的均匀分布,逻辑链路的带宽需求服从 [0, 50] 的均匀分布。每个请求的生命周期服从平均 1 000 个时间单位的指数分布,每次模拟仿真时间为 50 000 个时间单位。

3.2 实验设计

仿真实验共分为 4 组,对应跳数约束 h 取值为 2、3、4、5 时的场景。对于每组实验,首先构建

基于跳数约束的服务承载网拓扑 G^V ,记录并分析 G^V 的构建开销;然后随机生成 5 组服务请求,分别对应请求到达率为 4、5、6、7、8 的场景,用以测试服务供应质量在不同资源需求压力下的表现;最后对于每组请求,在 G^V 上运行 LG 算法构建服务路径,记录并计算性能评价指标。

3.3 实验结果与分析

图 2 为服务承载网的构建开销对比情况。图 2 表明,服务承载网的构建开销随着跳数约束值的增加而逐渐增加,主要原因在于链路数量的增加导致开销增加。当跳数约束达到某阈值时,服务承载网拓扑将成为完全图,构建开销收敛。图 3~7 显示了服务供应质量的主要评价指标在不同服务承载网和不同请求到达率下的表现。以请求到达率分别等于 4 和 8 时的情况为例,与跳数约束为 2 时相比,跳数约束为 3 时的请求接受率分别提高了约 7% 和 18%,平均收益分别提高了约 12.2% 和 51.8%,平均开销分别增加了 5.6% 和 40%;与跳数约束为 3 时相比,跳数约束为 5 时的请求接受率分别提高了约 0.1% 和 2%,平均收益分别提高了约 0.1% 和 2.8%,平均开销分别减少了约 1.9% 和 1.8%。当跳数约束为 2 时,服务承载网构建开销虽然最小,但是链路数量过少,在资源需求增长的压力下,承载能力不足的问题愈加严重,指标急剧下降,难以保证服务供应质量,不适合在请求到达率大于 5 的环境中使用;而跳数约束为 4 或 5 时,虽然链路数量充裕,承载能力增强,在各项评价指标中均表现较好,但是并未和跳数约束为 3 时的对应值拉开明显的差距,且构建开销分别增加了 50% 和 75%,性价比低。

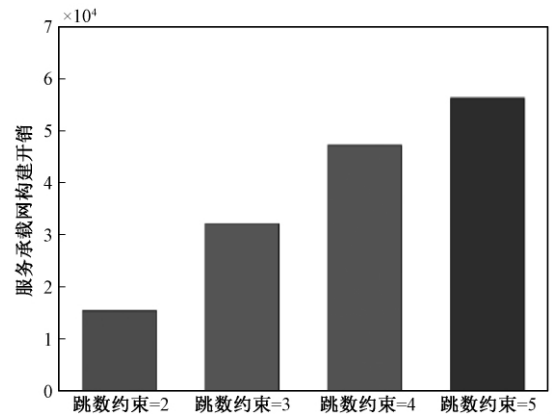


图 2 服务承载网的构建开销对比

Fig. 2 Comparison of cost of service carrying network

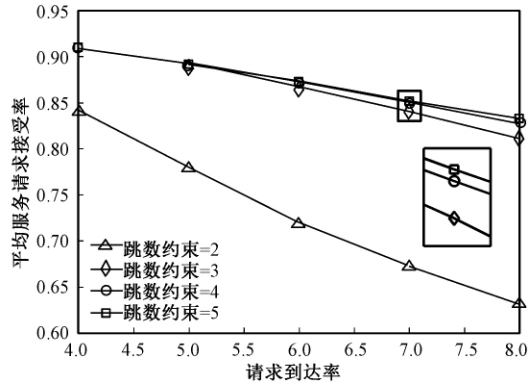


图3 平均服务请求接受率对比

Fig. 3 Comparison of average request acceptance ratio

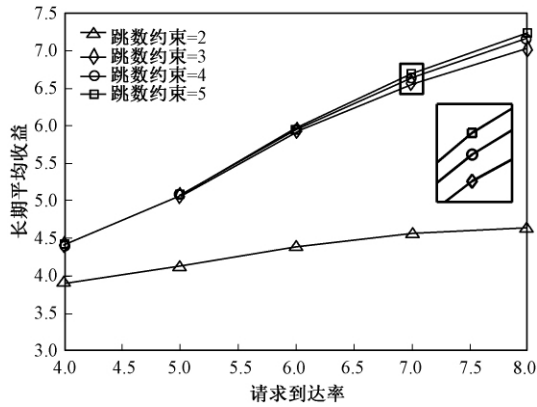


图4 服务承载网长期平均收益对比

Fig. 4 Comparison of long-term average revenue

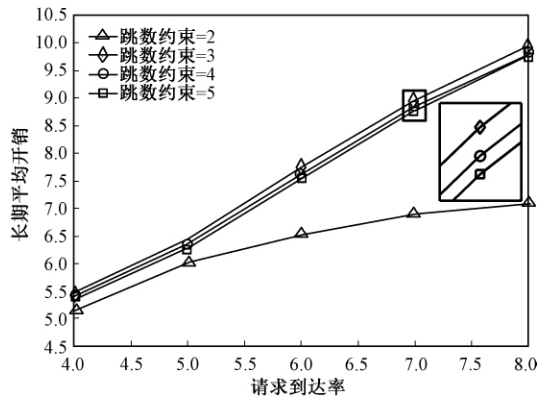


图5 服务承载网长期平均开销对比

Fig. 5 Comparison of long-term average cost

4 结论

笔者根据可重构网络的实际需要,提出了一种基于业务划分的新型端到端服务供应模型,根据请求所属的业务类别构建服务承载网,从而提供一致性的网络拓扑和资源视图.为验证所提出模型的可行性和有效性,设计了基于跳数约束的算法,在底层网络之上构建服务承载网,并使用 Layered Graph 算法在服务承载网之上构建服务路径.模拟实验结果验证了服务供应模型和所使

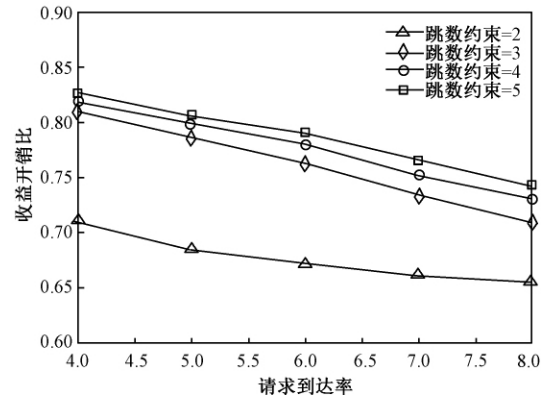


图6 服务承载网长期平均收益开销比对比

Fig. 6 Comparison of long-term revenue/cost ratio

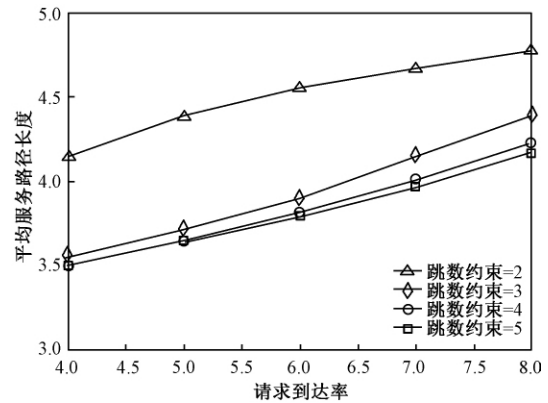


图7 平均服务路径长度对比

Fig. 7 Comparison of average length of service path

用算法的有效性,确定了性价比最优的跳数约束值.当跳数约束为3时,取得了与跳数约束为4或5时接近的性能,同时分别节省了约50%和75%的服务承载网构建开销.

如何在提出的服务承载网之上构建最优服务路径,对同类型业务流进行优化,是下一阶段研究工作中主要解决的问题.

参考文献:

- [1] CLARK D D, WROCLAWSKI J, SOLLINS K R, et al. Tussle in cyberspace: defining tomorrow's internet [J]. IEEE/ACM transactions on networking, 2005, 13(3): 462-475.
- [2] SALTZER J H, REED D P, CLARK D D. End-to-end arguments in system design [J]. ACM transactions on computer systems, 1984, 2(4): 277-288.
- [3] 兰巨龙, 程东年, 胡宇翔. 可重构信息通信基础网络体系研究 [J]. 通信学报, 2014, 35(1): 128-139.
- [4] 王子厚, 韩言妮, 林涛, 等. 可重构网络中基于中心度与拓扑势排序的资源分配算法 [J]. 通信学报, 2012, 33(8): 10-20.

- [5] HU Y X, LAN J L, WU J X. Providing personalized converged services based on flexible network reconfiguration [J]. Science China information sciences, 2011, 54(2): 334–347.
- [6] 梁宁宁, 兰巨龙, 程国振, 等. 基于拍卖博弈的可重构服务承载网动态构建算法 [J]. 电信科学, 2015, 31(5): 82–87.
- [7] CHOWDHURY M, RAHMAN M R, BOUTABA R. ViNEYard: virtual network embedding algorithms with coordinated node and link mapping [J]. IEEE/ACM transactions on networking, 2012, 20(1): 206–219.
- [8] HUANG X, GANAPATHY S, WOLF T. Evaluating algorithms for composable service placement in computer networks [C]//IEEE International Conference on Communications. Dresden: IEEE, 2009: 1–6.
- [9] TRAN K T, AGOULMINE N, IRAQI Y. Cost-effective complex service mapping in cloud infrastructures [C]//Proceedings of the IEEE Network Operations and Management Symposium. Hawaii: IEEE, 2012: 1–8.
- [10] 王宗江, 杨淑慧. 基于动态信任的最优云服务选择算法 [J]. 郑州大学学报(工学版), 2015, 36(4): 124–128.
- [11] MOENS H, TURCK F D. VNF P: A model for efficient placement of virtualized network functions [C]//International Conference on Network and Service Management. Rio de Janeiro: IEEE, 2014: 418–423.
- [12] LUIZELLI M C, BAYS L R, BURIOL L S, et al. Piecing together the NFV provisioning puzzle: efficient placement and chaining of virtual network functions [C]//IFIP/IEEE International Symposium on Integrated Network Management. Ottawa: IEEE, 2015: 98–106.
- [13] SAHHAF S, TAVERNIER W, ROST M, et al. Network service chaining with optimized network function embedding supporting service decompositions [J]. Computer networks, 2015, 93: 492–505.
- [14] ZEGURA E W, CALVERT K L, BHATTACHAEJEE S. How to model an internetwork [C]//Proceedings of the 15th Annual Joint Conference on IEEE Computer Communications Societies. San Francisco, CA, USA: IEEE, 1996: 594–602.

A Novel End-to-end Service Provisioning Model in Reconfigurable Network

MA Ding^{1,2}, ZHUANG Lei¹, LAN Julong³, MU Xiaowu⁴

(1. School of Information and Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China; 3. National Digital Switching System Engineering & Technological Research Center, Zhengzhou 450002, China; 4. School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Current service carrying network mapping algorithms ignored the business category of the service request, which resulted in the disconnection between the business carrying phase and the service provisioning phase. To address this problem, a novel end-to-end service provisioning model was proposed, in which service carrying networks were constructed considering aggregating service requests belonging to the same business category and customized end-to-end services were provided via establishing the corresponding service path over the service carrying network. To evaluate the feasibility and effectiveness of the proposed model, a service carrying network mapping algorithm with hop constraint was and designed Layered Graph algorithm was used to select service path. Simulation experiments on a large mix of service requests showed the impact the hop constraint had on network performance in terms of cost, revenue, acceptance ratio, length of service path, etc. under different working load, validated the effectiveness of the proposed model and algorithms, and eventually determined the optimal value of hops. When the value of hops equaled 3, compared with the situation in which the value of hops equals 4 or 5, the achieved performance was approximately same, and about 50 percent and 75 percent cost of constructing service carrying network can be saved respectively.

Key words: reconfigurable network; atomic capability; service carrying network; service provision model; service path

文章编号: 1671-6833(2017)06-0017-06

单云服务提供者环境下的随机化属性保护研究

李拴保^{1,2}

(1. 河南财政金融学院 信息工程系, 河南 郑州 451464; 2. 武汉大学 空天信息安全与可信计算教育部重点实验室, 湖北 武汉 430072)

摘 要: 单云服务提供者环境下用户随机属性隐私保护包括防范属性集更新泄露与密钥关联属性泄露, 主要通过代理认证、零知识证明、可信第三方和匿名签名实现. 针对属性保护严重依赖第三方的密钥分配与属性授权, 提出了一种密文策略属性基群签密随机属性保护方案. 该方案利用无证书群签密的无连接交互验证特性, 在用户计算密钥因子时防范云服务提供者获得密钥关联属性信息; 利用属性撤销的密钥重构与属性分割的密文重构的相互独立特性, 降低了用户签密所需的最小属性集数量, 抵制了攻击者利用属性集更新伪造签名; 以密钥服务为中心设计了群签密的身份验证机制以防范其他用户身份伪装. 结果表明: 该方案实现了保护随机属性安全和消息隐私.

关键词: 密文策略属性基加密; 密钥; 签名; 验证; 不可伪造

中图分类号: TP393.08

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.004

0 引言

云服务提供者 (cloud service provider, CSP)、授权者 (private key generation, PKG)、用户和数据属主 (owner) 密切合作, 才能有效实施大数据的处理与分析. 单 CSP 环境下, PKG 提供身份属性注册、密钥分配和授权服务, 存在用户隐私属性和密钥信息泄露. 因此, 隐私属性安全是单 CSP 平台的严重威胁之一, 现有方案主要是从匿名技术和零知识证明视角保障用户隐私属性安全.

(α, \mathbf{I})-匿名系统, 每个准身份 QI (quasi-identifiers) 至少由 \mathbf{I} 个记录共享, QI 是机密的, 其关联敏感属性值 s 不大于门限 α . QI 值和敏感值分为两个表格^[1], 元组分为多个桶^[2], 敏感值随机关联元组桶, 这样可以切断 QI 值和敏感值之间的关联. 针对存在隐私属性, Ercan 给出了 $Pr(t \in T | T^*) = \frac{m}{n}$ 的概率定义, 隐私属性存在泄露主要取决于 n 的规模. 类化技术^[3-4] 准身份值被具体值取代, 源数据集分成组, 每组至少包含 k 个元素, 存在信息损失.

零知识证明系统应用于单 CSP 环境身份认证; 个体属性集分布在第三方可信服务器^[5]; 多方计算协商云服务身份管理^[6], 这些方案均未涉及隐私保护. 云服务架构的身份理解和安全隐私限制^[7]、QI-属性不分割属性集模糊群签名算法^[8]、数据份额和群签名算法^[9]、自定义隐私保护策略^[10], 上述方法将身份泄露分为存在泄露和关联泄露. 存在泄露用于鉴别属性泄露, 关联泄露用于鉴别敏感属性泄露. 身份基群签密隐私保护^[11]和群撤销签名隐私保护^[12]方案, 均实现了部分身份属性保护和密文保护的细粒度访问控制. 但是, 可信 PKG 在系统中处于中心地位, 已经成为一个最大的系统瓶颈.

上述两种机制的共同特点是, PKG 轻负载、关联属性少, 身份保护与消息隐私保护相互独立, 这也是单 CSP 系统随机化属性保护的最大难点. 因此, 设计一种随机化属性保护方案对于单 CSP 系统的数据安全、密钥安全具有重要价值.

单 CSP 系统应用场景如图 1 所示, 用户隐私属性保护方案来源于扩展密文策略属性撤销^[13]方案与属性分割^[14]方案, 并且融合无证书群签密^[15]方

收稿日期: 2017-05-15; 修订日期: 2017-08-21

基金项目: 国家自然科学基金资助项目 (U1636107; 61373168), 河南省自然科学基金资助项目 (162300410191), 河南省软科学研究计划资助项目 (172400410501), 河南省科技攻关计划资助项目 (152102310245; 172102210172)

作者简介: 李拴保 (1972—), 男, 河南安阳人, 河南财政金融学院副教授, 博士, 主要从事大数据、云计算及信息安全方面的研究, E-mail: phdfuli@whu.edu.cn.

案. 用户、Authority 和 CSP 构成群, 用户代表群利用一组随机属性签密消息, Authority 鉴别其真实身份, 其他成员确信签密者来自群, 但不知道其具体信息, 保护了用户随机属性隐私, 用户可以访问云服务. 笔者提出的随机化属性基群签密的用户身份属性保护方案, 以密钥服务为中心, 引入扩展密文策略属性撤销的匿名性方法, 可以抵制用户之间共谋、CSP 与用户共谋获取身份属性信息; 引入属性分割的随机属性集代理密钥重构方法, 可以保护随机属性隐私和消息隐私, 结合无证书群签密的安全模型和密文公开验证方法, 达到抵制群成员身份伪装, 实现密文数据的细粒度访问控制的目的.

1 随机化属性保护系统基本框架

假设单 CSP 随机化属性保护系统存在用户之间共谋、CSP 与用户共谋及隐私泄露, 以图 1 应用场景为基础构造方案, 基本框架如图 2 所示. 具体步骤: 第①步, 用户向 PKG 申请属性授权和密钥分配, 获得私钥并生成签名上传 CSP; 第②步, CSP 和 PKG 协商用户的验证密钥, 并且验证用户签名, 若通过用户可访问云服务; 第③步, 用户获得 CSP 密文访问权. 群 O 基于实体成员的属性组成, 即 $O = \omega_1 \cup \omega_2 \cup \dots \cup \omega_N$, 其中, ω_i 为成员属性集. 群成员为 PKG、用户和 CSP, PKG 为群管理员.

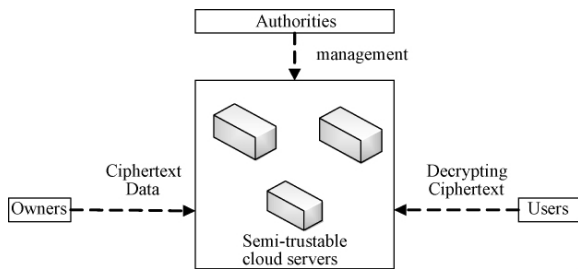


图 1 单 CSP 系统应用场景

Fig. 1 Single cloud service provider system application scenarios

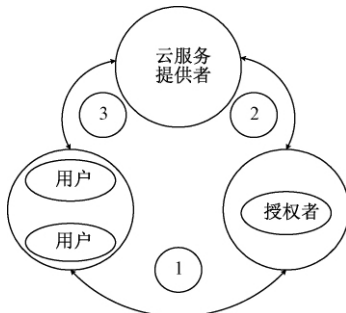


图 2 随机化属性保护系统具体运行流程

Fig. 2 Operation process of random attribute protection system

定义 1 密文策略属性基群签密随机化属性保护方案 (ciphertext-policy attribute based group signcryption with randomization attribute protection, CPAGSRAP) 是下列算法的一个元组.

Setup(λ): 给定安全参数 λ , 算法输出系统参数 $params$. PKG 选择随机整数 $n = p \times q$, 其中 p 和 q 为两个大素数; 选择 g, h 为 $GF(p)$ 的生成元. PKG 定义群 $O = \omega_1 \cup \omega_2 \cup \dots \cup \omega_N$, 选择字符集 ω_{PKG} 作为自身属性.

Partial-Private-key($params, \omega_{msk}$): PKG 选择随机整数 msk 作为主私钥, 计算公共参数 $mpk = g^{msk} \bmod n$.

Private-Key($params, mpk, msk, \omega_{PKG}$): 算法输入系统参数 $params$ 、公共参数 mpk 、主私钥 msk 、PKG 属性集 ω_{PKG} , 选择随机数 α 计算群公钥 G_{PK} 和群私钥 G_{SK} .

User-Key($params, \omega_U$): 算法输入系统参数 $params$ 、用户属性集 ω_U , 用户随机选择秘密值 w , 计算 $\omega_U = \omega_{PKG}^w$ 并发送给 PKG; PKG 选择秘密值 x , 计算基于 $(\omega_U, params, G_{SK}, x)$ 的 $(\sigma_1, \sigma_2, \sigma_3)$ 并发送给用户; 用户通过方程验证参数的真实性.

Signcryption($O, m, \omega_U, AS, params, \sigma_1, \sigma_2, \sigma_3, G_{PK}$): 算法输入消息 m 、群 O 、用户 ω_U 、访问结构 AS 、群公钥 G_{PK} 和 σ_3 , 用户选择秘密值 β 计算签密文 CT 发送给 CSP.

Delegate-PrivateKey($params, \gamma, mpk, msk$): 算法输入系统参数 $params$ 、更新属性集 γ ($\gamma \cap \omega_{PKG} = \emptyset$)、主私钥 msk 和公共参数 mpk , 计算新的群公钥 G'_{PK} 、群私钥 G'_{SK} 和代理密钥 G_{rk} .

ReSigncryption(CT, G_{rk}, β): 算法输入签密文 CT 、代理密钥 G_{rk} 、访问结构 AS 的一组属性集 β , 计算重签密文 CT 并发送给 CSP.

ReKey(G_{SK}, G_{rk}, θ): 算法输入群私钥 G_{SK} 、代理密钥 G_{rk} 、 G_{rk} 与 G_{SK} 的共同属性集 θ , 输出用户更新私钥 SK' .

Unsigncryption(O, CT, G_{PK}): 算法由 CSP 运行, 算法输入签密文 CT 、群 O 和群公钥 G_{PK} , 验证 CT 有效且签密者系 O 成员, 输出 "valid" 或 "Invalid".

Verify(CT, G_{SK}): 算法由 PKG 运行, 输入签密文 CT 、群私钥 G_{SK} , 验证签密文 CT 的正确性.

定义 2 一个密文策略属性基群签密随机化属性保护方案 (CPAGSRAP) 在自适应选择密文攻击下是安全的, 条件是所有的多项式时间算法攻击者在 IND-CPAGSRAP-CCA2 游戏模型中获胜的

概率最多具备一个可忽略的优势。

2 密文策略属性基群签密随机化属性保护方案

2.1 初始化阶段

系统参数、密钥生成环境的初始化,为用户访问云服务提供认证准备. 系统定义多项式群上的拉格朗日插值公式 $\Delta_{i,S}$,用于密钥分配管理; 对任意 $i \in \mathbf{Z}_p$, 假设 S 是 \mathbf{Z}_p 中的 d -元素集合, 则

$$\Delta_{i,S(x)} = \prod_{j \in S, j \neq i} \frac{x - j}{i - j}. \quad (1)$$

系统定义群 $O = \omega_1 \cup \omega_2 \cup \dots \cup \omega_N$; PKG 选择自身的属性集 $\omega_{\text{PKG}} \subset O$. 给定安全参数 λ , PKG 选择 G 加法循环群、 G_T 乘法循环群, 两个阶均为素数 p ; 双线性映射 $e: G \times G \rightarrow G_T, g, h$ 是 G 的一个生成元; 系统选择整数 n 且满足 $n = p' \times q'$, 其中 $p' = 2p'_1q'_1 + 1$ 和 $q' = 2p'_2q'_2 + 1$, 以及 $p', q', p'_1, q'_1, p'_2, q'_2$ 均为大素数. 假设 U 是通用属性的集合, 且 $|U| = L$; $\Omega = \{\Omega_1, \dots, \Omega_{d-1}\}$ 是一个 $d-1$ 缺省属性集合, 满足拉格朗日插值公式 (1).

系统定义密码学哈希函数 $H: \{0,1\}^* \rightarrow \{0,1\}^{|n|}$, $|n|$ 表示签密文的长度, 用于抵制用户之间共谋. 系统选择随机数 $\alpha, \beta \in \mathbf{Z}_p$, 计算 $Y = e(g, g)^\alpha$. 系统构造属性树 \mathbf{T} 是由群 O 属性集构成的访问结构, T_x 是某一节点 x 的 \mathbf{T} 的子树, 用户随机属性集 ρ 满足 T_x 即 $T_x(\rho) = 1$, 当且仅当 ρ 满足访问策略 π 即 $T_\pi(\rho)$ 是真实用户可验证签名.

系统选择随机数 $t_i \in \mathbf{Z}_p (1 \leq i \leq 3n)$, 计算 $T_i = g^{t_i}$, 系统生成 $PK = (e, g, Y, T_1, \dots, T_{3n})$ 和 $MK = (\alpha, t_1, \dots, t_{3n})$. PKG 发布公共参数 $params = \{e, O, G, G_T, n, g, H, Y, \pi, h, PK, MK\}$, 保存私有参数 (α, β, t_i) .

2.2 密钥服务阶段

2.2.1 部分私钥生成

PKG 为用户生成部分密钥因子, 用户利用部分参数与 PKG 交互认证, 实现本地密钥的生成.

系统定义用户随机属性集 $\omega_u' \subset O$ 并且计算 $H(\omega_u', O)$, 随机选取一个属性子集 $\omega_U \subset \omega_u'$ 并且计算 $H(\omega_U, O)$; 用户通过安全通道向 PKG 发送元组 $(\omega_U, H(\omega_u', O), H(\omega_U, O))$.

PKG 验证 $\omega_U \subset G1$, 可以确认用户是群成员, PKG 随机选择私有整数 msk , 计算公共参数 $mpk = g^{msk} \bmod n$, 假设 $\hat{\omega}_U = \omega_U \cup \Omega$.

PKG 选择随机数 $\gamma_s \in \mathbf{Z}_p$, 设用户签名属性集

$\rho_s \subset \hat{\omega}_U$ 且满足 $T_\pi(\rho_s)$ 为真, 系统计算用户签名私钥部分因子 $D_s = \llbracket (\alpha + \beta) / \gamma_s \rrbracket h$.

PKG 选择随机数 $\gamma_E \in \mathbf{Z}_p$, 设用户加密属性集 $\rho_E \subset \hat{\omega}_U$ 且 $\rho_s \cap \rho_E = \Phi$ 且满足 $T_\pi(\rho_E)$ 为真, 系统计算用户加密私钥部分因子 $D_E = \llbracket (\alpha + \beta) / \gamma_E \rrbracket h$.

2.2.2 私钥生成

PKG 为用户定义 CSP 成员访问入口, PKG 选择随机数 $v \in \mathbf{Z}_p$, 系统计算 $D'_{SK} = v \cdot mpk \cdot D_s + H(\omega_{\text{PKG}}) \cdot H(\omega_U, O) \bmod n$, 其中 D'_{SK} 为群私钥; $D'_{PK} = g^{D'_{SK} + D_E} \bmod n$, 其中 D'_{PK} 为群公钥.

RSA 算法 $d \cdot e = 1 \bmod n$, e 是公钥, d 是私钥. 系统定义参数 $(n, g, D'_{PK}, e, H(\omega_{\text{PKG}}))$ 为群公钥, 参数 (d, D'_{SK}) 为群私钥.

用户选择随机数 $w \in \mathbf{Z}_p$, 系统计算 $H(\hat{\omega}_U) = H(\omega_{\text{PKG}})^w \bmod n$, 发送给 PKG. PKG 计算 $M_U = (H(\hat{\omega}_U))^d, (M_U, d)$ 为用户访问 CSP 的成员入口.

2.2.3 用户密钥生成

用户获得系统生成的群公钥和群私钥, 系统为用户生成群签名密钥和验证密钥. PKG 选择随机数 $r_i \in \mathbf{Z}_p, i \in U$. 设 $r = \sum_{i=1}^n r_i$, 用户计算 $SK = (D, \bar{D} = \{D_i, F_i\}_{i \in U})$, 其中 $D = g^{\alpha - r}$. 如果 $i \in U$, 那么 $D_i = g^{\frac{r_i}{t_i}}$ 且 $F_i = g^{\frac{r_i}{t_{2n+i}}}$; 如果 $i \notin U$, 那么 $D_i = g^{\frac{r_i}{t_{n+i}}}$.

用户签名私钥为 $D''_{SK} = D^{D_{SK}} + rD_i$, 验证密钥为 $D''_{PK} = r F_i^{D_{PK}}$.

2.3 签密与验证服务阶段

2.3.1 用户群签密

用户获得访问 CSP 的入口, 向 CSP 发送签名消息和加密消息. CSP 验证用户系群 $G1$ 的真实成员, 系统计算签密文.

系统选择随机数 t , 定义签名属性子集 $\rho_s \subset \hat{\omega}_U$, 计算 $H(\rho_s, O)$, 定义加密属性子集 $\rho_E \subset \hat{\omega}_U$ 且满足 $\rho_s \cap \rho_E = \Phi$, 计算 $H(\rho_E, O)$; 系统定义签名策略 π_s 和密文策略 π_E , 定义访问控制树 T_s 和 T_E ; 系统计算: $C_1 = g^t, C_2 = m \oplus Y^t$.

系统选择随机数 k , 系统计算: $V = e(C_1, h)^k$, $C = H(m, C_2, V, \pi_s, \pi_E), T = [C_2]g^e + [C^d]D_s, B_{s,E} = [T, H(\rho_s, G1)]D''_{SK} + [C, H(\rho_E, O)]D_E$.

系统选择随机数 α', β' , 系统计算:

$$\delta_1 = (H(\hat{\omega}_U))^{\alpha'} \bmod n; \delta_2 = \alpha' \cdot T + \delta_1 \bmod n;$$

$$\delta_3 = (M_U)^{\delta_1 \cdot e} \bmod n; C' = \beta' \cdot M_U + D''_{PK} \bmod n.$$

系统计算 $M_U^{\delta_2} = \delta_1^T \cdot \delta_3 \bmod n$ 成立, 系统输

出签密文 $CT = (T_s, C_1, C_2, B_{s,\omega}, T_E, C, T, \delta_1, \delta_2, \delta_3, C)$.

2.3.2 重代理密钥生成

用户委托中间用户生成代理密钥. 设 $i \in \gamma$, γ 满足 $[1, 2n]$, i 的值小于等于 n 为有效属性; 反之, 大于等于 n 时有效值为 $i - n$.

对任意 $i \in \gamma$, 选择随机数 t'_i , 中间用户计算 $rk_i = \frac{t'_i}{t_i}$, 对每一个 $i \in \{1, 2, \dots, 2n\} / \gamma, rk_i = 1$.

用户重代理密钥 $G_{rk_i} = \{rk_i\}_{1 \leq i \leq 2n}$. 系统定义新的群公钥 $D''_{PK} = G_{rk_i} D''_{PK}$, 新的群私钥 $D''_{SK} = G_{rk_i} D''_{SK}$.

2.3.3 重签密

系统定义签密文的访问结构 $AS = \bigwedge_{i \in I} \tilde{t}_i, \beta \in [1, 2n]$. 重签密签密文算法如下:

对任意 $i \in \beta, C'_i = C_i^{G_{rk_i}}$ 且 $1 \leq i \leq n; n < i \leq 2n, C'_{i-n} = C_i^{G_{rk_i}}$. 对任意 $i \in U, C'_i = C_i$ 且 $i \notin \beta$ 和 $i + n \notin \beta$.

系统计算重签密文 $CT' = (AS, C_1, C_2, C', \{C'_i\}_{i \in U})$ 并发送给签密 CSP.

2.3.4 重密钥生成

系统定义 $\theta \in [1, 2n]$, 对任意 $i \in \theta, D'_i = D_i^{G_{rk_i}^{-1}}$ 且 $1 \leq i \leq n$; 如果 $n \leq i \leq 2n, D'_{i-n} = D_i^{G_{rk_i}^{-1}}$.

对任意 $i \in U, D'_i = D_i$ 且 $i \notin \theta$ 和 $i + n \notin \theta$, 系统计算用户重密钥 $\bar{D}' = \{D'_i, F_i\}_{i \in U}$.

2.3.5 解签密与验证

CSP 验证重签密文 CT' 有效性以及用户身份属性的真实性, 可以证实用户为群 O 成员.

如果用户属性为 T_E 叶子节点, 系统完成计算

$$\text{和验证: } S_E = \frac{e(D'_i, D''_{PK})}{e(\delta_1, C_1)} = e(g, h)^{H(\omega_E, O)}.$$

如果用户属性为 T_S 叶子节点, 系统完成计算

$$\text{和验证: } S_S = \frac{e(h, D''_{PK})}{e(\delta_2, C_2)} = e(g, h)^{H(\omega_S, O)}.$$

用户一部分属性为 T_E 叶子节点, 另一部分属性为 T_S 叶子节点, 系统完成计算和验证:

$$S_{E,S} = \frac{e(C', D''_{PK})}{e(\delta_3, C)} \cdot e(g, h)^V = e(g, h)^{\alpha\beta},$$

$m = C_2 \oplus (S_{E,S})^C$, 并且 T_E 和 T_S 满足访问策略 $T_\pi(\rho)$.

3 性能分析

设 $|p|$ 表示 Z_p 的元素规模, $|g|, |g_T|$ 分别表示 G, G_T 的元素规模, n_a, n_u 分别表示系统用户属性数、用户总数, H_a 表示哈希函数计算, P_a 表示双线性对计算, Exp_G 表示 G 上的指数运算, Exp_{G_T} 表示 G_T 上的指数运算, $|aG|$ 表示 G 上的 a 元素二进制长度.

文献 [14] 的通信成本是系统参数、用户私钥, 私钥规模与 $|p|$ 相关; 文献 [15] 的通信成本是系统参数、主密钥和签名私钥. 本文方案、CSP 和用户之间的通信成本主要来自于签密文验证. 文献 [14 - 15] 的通信成本包括签名和属性集更新树, 与属性数量成线性关系; 签名规模与 $|g_T|$ 相关, 属性集规模与 $|g|, |p|$ 和 n_a 相关. 本文方案 PKG 的通信成本小于文献 [14 - 15], CSP 成本大于文献 [14 - 15], 比较结果如表 1 所示.

PKG 存储开销包含系统参数、主密钥和群密钥、群私钥, 系统参数与 n_u 成线性关系; 用户和 CSP 存储开销包含群密钥, 与 $|g|$ 和 $|g_T|$ 相关. 文献 [14] 中 PKG 存储开销包含属性集、系统参数、主密钥, 系统参数规模与 n_u 成线性关系; CSP 和用户存储开销主要是秘密值、密名、密文, 密文与 n_u 成线性关系; 文献 [15] 中 PKG 还增加了撤销属性集, 与 $|g|$ 相关. 本文方案 PKG 存储开销低于文献 [14 - 15], 用户开销相当, CSP 存储开销高于文献 [14 - 15], 比较结果如表 2 所示.

表 1 随机属性保护通信成本比较

Tab. 1 Communication cost comparison of random attribute protection

实体	本文方案	文献 [15]	文献 [14]
用户和 PKG	$2 g $	$ g + n_a g $	$2 g + 2n_a g $
CSP 和 PKG	$ g + g_T $	$ g + 3 p $	$ g + 3 p $
用户和 CSP	$ p + g_T + n_a g $	$ g_T + 2 p + g $	$2 g_T + n_u p $

表 2 随机属性保护存储开销比较

Tab. 2 Storage cost comparison of random attribute protection

实体	本文方案	文献 [15]	文献 [14]
PKG	$n_a p $	$n_u g + 2 p $	$2n_u g + p $
用户	$(4 + n_a) g $	$2n_u g + 5 p $	$3n_u p + g $
CSP	$n_u g + n_u g_T $	$3 g_T + g $	$3 g_T + 3n_u g $

系统计算效率主要包含计算时间和计算成本,群签密与解签密验证计算时间主要与用户属性数量相关,计算成本与困难性假设相关. 用户属性数量规模增加,系统计算时间增长率决定了签密和验证的效率,通过仿真实验方法测试方案的签密和验证计算时间增长率. 仿真实验结果如图3所示,纵轴表示系统计算时间,横轴表示用户属性数量. 本文方案密文策略群签密与解签密验证服务系统在用户属性数量为(20,80)时计算时间增长率小于文献[14-15]. 在双线性映射和标准模型下,密文策略属性基群签密部分属性保护方案在计算成本方面有较大改进. 与文献[14-15]相比,本文方案在指数运算、哈希运算方面优于文献[14-15],密文规模运算高于文献[14-15].

在 DBDH 困难性假设适应性选择密文攻击和选择消息攻击下,方案存在签名强不可伪造性、身份属性匿名性,比较结果如表3所示.

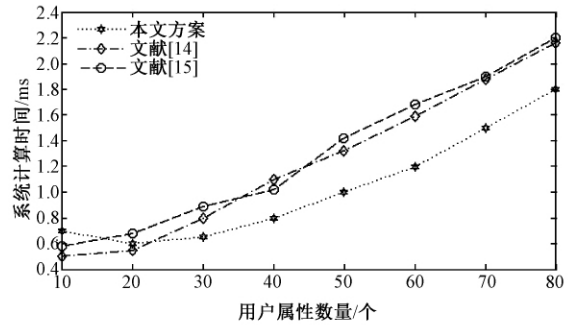


图3 随机属性保护计算时间比较

Fig.3 Calculated time comparison of random attribute protection

表3 随机属性保护计算成本比较

Tab.3 Calculated costs comparison of random attribute protection

计算成本指标	本文方案	文献[15]	文献[14]
对运算	$2P_a$	$3P_a$	$4P_a$
指数运算	$Exp_G + Exp_{G_T}$	$4Exp_G + Exp_{G_T}$	$3Exp_G + 3Exp_{G_T}$
哈希计算	$2H_a$	$6H_a$	$3H_a$
密文规模	$ G + G_T $	$3 G + 3 G_T $	$2 G + 2 G_T $
采用模型	标准	RO	RO

综合单 CSP 系统通信成本、存储开销和计算效率3个方面,本文方案在 PKG 通信成本和存储开销以及签密验证计算时间优于文献[14-15],但是密文规模、CSP 通信成本和存储开销仍然需要进一步研究和优化.

4 结论

身份属性安全是单 CSP 系统安全的焦点,密文策略属性基群签密部分属性保护方案,以双线性映射、DBDH 假设标准模型下的无证书签密为基础,对其匿名认证、密钥管理进行了群签密方案相关扩展,融合了密文策略属性基密码机制,解决了用户部分身份属性泄露、身份属性伪造问题. 用户在申请密钥服务之前降低了和 PKG 的交互通信;群签密与验证的方法简化了身份认证的计算复杂度. CPAGSRAP 方案简化了群签密的指数运算次数,消除了证书存储负载,降低了群签密算法和解签密算法的运算负载. 但增加了群密钥与签密验证的运算次数和 CSP 综合开销,因此如何降低综合负载将作为进一步的研究方向.

参考文献:

- [1] XIAO X, TAO Y. Anatomy: simple and effective privacy preservation [C] // Proc of Very Large Data Base Conference. Seoul: Spring, 2006: 139 - 150.
- [2] ZHANG Q, KOUDAS N, SRIVASTAVA D, et al. Aggregate query answering on anonymized tables [C] // Proc of International Conference on Data Engineering. Istanbul: Spring, 2007: 116 - 125.
- [3] BAYARDO R J, AGRAWAL R. Data privacy through optimal k-anonymization [C] // Proc of International Conference on Data Engineering. Tokyo: Spring, 2005: 217 - 228.
- [4] MEYERSON A, WILLIAMS R. On the complexity of optimal k-anonymity [C] // Proc of ACM International Conference on Principles of Database Systems. Paris: Spring, 2004: 223 - 228.
- [5] HUSSAIN M. The design and applications of a privacy-preserving identity and trust-management system [D]. Kingston, Ontario, Canada: School of Computing, Queen's University, 2010.
- [6] RANCHAL R, BHARGAVA B K, OTHMANE. Protection of identity information in cloud computing without trusted third party [C] // Proc of IEEE Symposium on Reliable Distributed Systems. Paris: IEEE, 2010: 368 - 372.
- [7] HARALAMBOS M, SHAREEFUL I. A framework to support selection of cloud providers based on security

[1] XIAO X, TAO Y. Anatomy: simple and effective pri-

- and privacy requirements [J]. Journal of systems and software, 2013, 86(6): 2276–2293.
- [8] WANG H. Privacy-preserving data sharing in cloud computing [J]. Journal of computer science and technology, 2010, 25(3): 401–414.
- [9] CHUANG I H, LI S H, HUANG K C, et al. An effective privacy protection scheme for cloud computing [C]//2011 13th International Conference on Advanced Communication Technology(ICACT) . Seoul, Korea: IEEE, 2011: 13–16.
- [10] CHADWICK D W, FATEMA K. A privacy preserving authorisation system for the cloud [J]. Journal of computer and system sciences, 2012, 78(10): 1359–1373.
- [11] QIN L, GUO J W. Time-based proxy reencryption scheme for secure data sharing in a cloud environment [J]. Information sciences, 2015, 258(10): 355–370.
- [12] XUE R W, JIAN M F. User key revocation method for multi-cloud service providers [J]. Journal of electronics information technology, 2015, 37(9): 2225–2231.
- [13] YU S, WANG C, REN K, et al. Attribute based data sharing with attribute revocation [C]//ASIACCS' 10 Proceeding of the 5th ACM Symposium on Information, Computer and Communications Security. New York: ACM, 2010: 261–270.
- [14] LIF W, HAO J Z. Security and privacy for storage and computation in cloud computing [J]. Information sciences, 2015, 258(4): 371–386.
- [15] MOHANTY S, MAJHI B, DAS S. A secure electronic cash based on a certificateless group signcryption scheme [J]. Mathematical and computer modelling, 2013, 58(1/2): 186–195.

Research on Randomization Attribute Protection in Single Cloud Service Provider

LI Shuanbao^{1,2}

(1. Department of Information Engineering, Henan College of Finance, Zhengzhou 451464, China; 2. Key Lab of Aerospace Information Security and Trusted Computing Ministry of Education, Wuhan University, Wuhan 430072, China)

Abstract: User randomization attribute privacy protection included attribute set updating leakage and attribute leakage of key in Single-CSP (Cloud Service Provider), which performed mainly through proxy authentication, zero-knowledge proof, the trusted third party and anonymous signature. Focusing on attribute protection heavily dependent on third-party key distribution and attribute authority, this paper presented a ciphertext policy attribute-based group signcryption randomization attribute protection scheme. When a user calculated the key factor, the scheme controlled CSP getting key associated attribute information by using certificateless group signcryption connectless cross-validation; it reduced the minimal number of attribute set for signcryption need by using attribute revocation and attribute segmentation to mutual independent with sign key and encryption key, and resisted an attacker forged signatures of attribute set updating. In key service-centric, it designed the identity verifying mechanism of group signcryption, and controls masquerading as other user. The scheme implemented the protection of the randomization attribute security and message privacy.

Key words: CP-ABE (ciphertext policy attribute-based encryption); key; signature; verify; unforgeability

基于喷泉码及 CDMA 扩频的数字水印盲算法

张卫党, 吴琳

(郑州大学 信息工程学院, 河南 郑州 450001)

摘要: 为了提高数字图像水印系统的鲁棒性和水印容量, 提出一种基于喷泉码及 CDMA 扩频的数字图像水印算法. 该算法首先对原始水印信息进行 LT 码预编码和 CDMA 扩频处理, 然后将编码后的水印信息嵌入到载体图像的离散余弦变换 DCT (Discrete Cosine Transform) 域; 在水印检测时, 利用正交扩频码的互相关特性实现水印的盲检测, 并且设定判决阈值将出错的水印比特删除, 利用正确的水印比特进行 LT 码译码, 提高了水印的鲁棒性. 实验结果表明, 该算法既能够保证水印系统具有良好的不可见性, 又具有较大的水印容量, 且对常见的水印攻击有较好的鲁棒性, 尤其对 JPEG 压缩、噪声和剪切攻击有很强的鲁棒性.

关键词: 数字水印; CDMA; 喷泉码; DCT; 鲁棒性

中图分类号: TU528.1

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.005

0 引言

数字水印技术是一种信息隐藏技术, 其基本思想是在数字图像、视频和音频等数字载体中嵌入水印信息, 以达到版权保护、产品认证和信息隐藏等目的. 数字水印技术可以看作一个通信系统, 水印信息在传输过程中必然会受到有意或无意的攻击. 很多通信技术被引入到数字水印技术中以提高水印系统的可靠性等性能. 目前扩频 (spread spectrum) 技术已广泛应用到数字水印技术中, Cox 等^[1] 首先提出直接序列扩频水印的思想, 提高了水印系统的鲁棒性, 但直接扩频水印容量较低. 在通信系统中, 码分多址 CDMA (code division multiple access) 是一种有效的通信方式, 具有保密性好、容量大、可多址复用、抗噪声能力强等优点^[2-4]. 许多研究者已经将 CDMA 扩频技术用到数字水印系统中, Silvestre 等^[5] 给出了 CDMA 扩频水印在载体图像 DCT 变换域的数字水印方案; 文献 [6] 提出了基于 Gold 码的 CDMA 扩频数字水印算法, 仿真结果表明, CDMA 扩频方案比直接序列扩频方法具有更好的鲁棒性; 文献 [7] 给出了 CDMA 扩频与 Turbo 码在载体图像 DCT 和 DWT (discrete wavelet transform) 混合变换域的数

字水印方案, 实验结果表明, 应用 CDMA 扩频技术不仅大大提高了水印的鲁棒性, 也使得水印容量大大增加; 文献 [8] 分析了 CDMA 扩频水印系统信息嵌入量、扩频码长与水印误码率之间的关系.

目前, 很多研究者在数字水印系统中应用通信系统中的信道编码技术对水印信息进行预编码, 以提高水印的鲁棒性. 数字喷泉码^[9-10] (digital fountain code) 是一种稀疏矩阵码, 在删除信道下具有优越的性能. 编码的发送端是由 k 个原始数据生成的任意数量的编码分组, 在接收端只要能够接收到任意 $k(1 + \epsilon)$ 个编码分组, 即可实现完全译码, ϵ 是译码开销, 优良的喷泉码具有很小的译码开销 ϵ . 由于数字喷泉码的编译码方法简单, 采用其对原始水印信息进行预编码不仅提高了水印安全性, 而且可以大大提高水印的鲁棒性. LT 码^[11-13] 是 Michael Luby 在 2002 年提出的第一种实用的喷泉码, 文献 [14] 首先将 LT 码应用于数字水印系统, 实验结果表明, 采用 LT 码预编码的水印方案具有很强的抗剪切攻击能力; 文献 [15] 给出了 LT 码和双伪随机序列扩频结合的水印方案, 算法的鲁棒性得到了较好的改善.

基于 LT 码较低的编译码复杂度和良好的纠

收稿日期: 2017-03-26; 修订日期: 2017-06-28

基金项目: 国家自然科学基金资助项目 (U1204607)

作者简介: 张卫党 (1958—), 男, 河南郑州人, 郑州大学教授, 博士, 主要从事信道编码研究, E-mail: iewzhang@zzu.edu.cn.

删性能以及 CDMA 扩频具有的抗干扰、容量大等特性,笔者提出一种新的基于 LT 码和 CDMA 扩频的载体图像 DCT 变换域数字图像水印盲算法. 实验结果表明,该算法不仅具有较好的抗 JPEG 压缩能力、抗剪切攻击能力和抗噪声攻击能力,而且安全隐蔽性好,在大水印容量情况下也能完整地提取出水印信息.

第 1 节阐述了基于 LT 码编码和 CDMA 扩频处理的数字水印的生成方案;第 2 节给出了水印在载体图像 DCT 变换域的嵌入算法;第 3 节提出了水印的检测和提取方案;第 4 节给出了仿真实验的结果和分析,最后总结全文.

1 数字水印生成

1.1 LT 码预编码

首先将原始水印图像转换成长度为 r 的一维二进制序列 $z' = [z'_1, z'_2, \dots, z'_r]^T$, 然后对序列 z' 进行 LT 码编码, LT 码的度分布函数采用鲁棒孤子度分布, 其参数 $c = 0.03, \delta = 0.2$. 设经 LT 码编码后的水印信息长度为 L , 构造出 LT 码的生成矩阵 T , 计算 $z = Tz'$, 即得到预编码后的二进制水印信息序列 $z = [z_1, z_2, \dots, z_L]^T$. 为平衡下节 CDMA 扩频处理后的混合信号, 对水印序列 $z = [z_1, z_2, \dots, z_L]^T$ 进行由 $\{0, 1\}$ 到 $\{-1, 1\}$ 的映射, 即水印信息为“1”时映射为其本身, 水印信息为“0”时映射为“-1”, 则得到二值水印信息序列 $B' = \{b'(j) | j = 1, 2, \dots, L\}$, 其中 $b'(j) \in \{-1, 1\}$.

1.2 CDMA 扩频水印生成

在 CDMA 系统中, 扩频码要有良好的自相关和互相关特性, m 序列、Gold 码等虽然具有很好的自相关和互相关特性, 但 Gold 码的互相关性并不理想, 而 m 序列可用的地址码组数少. Walsh 码在同步时具有理想的互相关特性, 在数字水印系统中不考虑“用户”不同步问题, 因此, 笔者采用 Walsh 码作为扩频码. Walsh 码可由 Hadamard 矩阵生成, 依据密钥从 Hadamard 矩阵中选择 K 行组成 Walsh 码扩频码集 $G = \{g_k(n) | k = 1, 2, \dots, K; n = 1, 2, \dots, N\}$. 根据扩频码长度和可嵌入水印信息的载体图像 DCT 变换域系数的个数, 将上节得到的二值序列 B' 进行分组, 设共分为 K 组, 则每组有 $R = L/K$ 个元素, 则经 LT 码编码后水印信息可以表示为

$$B = \{b_k(i) | k = 1, 2, \dots, K; i = 1, 2, \dots, R\}, \quad (1)$$

其中, $b_k(i) \in \{1, -1\}$ 表示第 k 个用户传输的第 i 个水印比特. 应用 CDMA 扩频方法对水印信息扩频编码, 最终生成的待嵌入水印信息表示为

$$w_i = \sum_{k=1}^K b_{k,i} g_k, \quad i = 1, 2, \dots, R, \quad (2)$$

式中: g_k 为集合 G 的第 k 个元素; w_i 为 N 维向量.

2 水印嵌入算法

离散余弦变换是一种正交实变换, 计算量适中且具有较好的去相关压缩能力, 同时与压缩标准兼容, 因此, 本文水印信息选择嵌入到载体图像的 DCT 变换域的中频系数中. 设载体图像为 I , 将其进行 8×8 分块, 对每个 8×8 的图像子块进行 DCT 变换, 依据密钥选取每个图像子块的 DCT 变换域的某些中频系数, 将待嵌入水印信息依次嵌入, 水印嵌入采用常用的加性嵌入规则, 如式 (3) 所示.

$$X_i^w = X + \alpha w_i, \quad i = 1, 2, \dots, R, \quad (3)$$

式中: X_i^w 为嵌入水印后的载体图像 DCT 变换域系数集; α 为嵌入强度; X 为嵌入水印前的载体图像系数集; w_i 为上节生成的待嵌入水印信息.

最后再对嵌入水印后的每个子图像块进行逆离散余弦变换, 从而生成嵌有水印的载体图像 I_w .

3 水印的检测和提取算法

水印检测是利用扩频码的线性相关特性, 不需要原始载体图像和原始水印图像参与, 实现了水印盲检测. 首先将接收到的嵌有水印信息的载体图像 I_w 进行 8×8 分块, 再对每个图像块进行 DCT 变换, 根据嵌入时的密钥依次取出嵌有水印的 DCT 变换域中频系数, 并且按照嵌入前水印信息的顺序构造出含有水印信息的系数矩阵, 则水印检测端输入为

$$X_i^{wn} = X_i^w + n_0, \quad i = 1, 2, \dots, R. \quad (4)$$

然后在 N 维 Hadamard 矩阵中, 根据密钥选择出与水印信息扩频时相同的 Walsh 码正交序列集 G 用于解扩. 由于 Walsh 扩频码具有正交的特性, 即不同码间的多址干扰为零. 据此, 将检测器输入端的载体图像 DCT 变换域系数集 X_i^{wn} 与扩频码 g_k 作线性相关, 其结果为

$$\begin{aligned} \eta_{k,i} &= \frac{1}{N} ((X_i^{wn}) \cdot g_k) = \\ &= \frac{1}{N} ((X + \alpha w_i + n_0) \cdot g_k) = \\ &= \frac{1}{N} (X \cdot g_k + \alpha \sum_{j=1}^K b_{j,i} g_j \cdot g_k + n_0 \cdot g_k). \end{aligned} \quad (5)$$

由于 $\{g_k\}$ 是正交的,即当且仅当 $j = k$ 时, $g_j \cdot g_k$ 为非零值,则上式可以写为

$$\eta_{k,i} = \frac{1}{N}(\mathbf{X} \cdot \mathbf{g}_k) + \frac{1}{N}(n_0 \cdot \mathbf{g}_k) + \alpha b_{k,i}, (6)$$

式中: $\eta_{k,i}$ 为检测出的第 k 个用户的第 i 个水印比特的相关值,它反映了提取出的 DCT 变换域中频系数与嵌入的水印信息的相关程度. 由于载体图像 DCT 变换域系数和噪声与正交扩频码无关,使得相关检测值 $\eta_{k,i}$ 的正负与水印比特 $b_{k,i}$ 保持一致. 由于水印信息在传输过程中会受到各种攻击,因此在相关检测时设置判决阈值 τ ,当相关值 $\eta_{k,i}$ 超过阈值 τ 时判定该系数中含有水印信息,并且根据 $\eta_{k,i}$ 的正负判定嵌入的水印比特是“1”或者“-1”;否则,该处视为未检测出水印信息,并把 LT 码的编码生成矩阵 \mathbf{T} 中的相应列删除. 据此对相关检测结果进行如下判决:

$$\tilde{b}_{k,i} = \begin{cases} 1, & \|\eta_{k,i}\| > \tau \ \& \ \eta_{k,i} > 0; \\ 0, & \|\eta_{k,i}\| > \tau \ \& \ \eta_{k,i} < 0, \end{cases} (7)$$

其中, $\tilde{b}_{k,i}$ 为判决得到的第 k 个用户的第 i 个水印比特; τ 为判决阈值,它依据实验确定. 在水印嵌入时为平衡 CDMA 扩频后的信号,对 LT 码编码后的水印信息进行了“0”到“-1”的映射. 因此在判决式(7)中,当相关值 $\eta_{k,i} < 0$ 时直接判决此时的水印比特为“0”.

最终,由提取出的信息 $\tilde{b}_{k,i}$ 及更新过的生成矩阵 \mathbf{T} 进行 LT 码译码即可得到原始水印信息的估值序列 $\hat{\mathbf{b}} = [\hat{b}_1, \hat{b}_2, \dots, \hat{b}_r]^T$.

4 实验结果和分析

实验使用 512×512 的 lena 灰度图像作为原始载体图像,水印信息采用带“数字水印”文字标识的 32×32 的二值图像,分别如图 1(a)、(b)所示. 原始载体图像共可以划分成 4 096 个 8×8 子块,编码后的水印信息依据密钥依次嵌入到 4 096 个载体图像子块的 DCT 变换域中频系数中,用峰值信噪比(R_{PSN})来衡量嵌入水印后载体图像的失真情况.

实验原始水印信息的 LT 码编码长度 $L = 4\ 096$,为方便对其进行 CDMA 扩频处理,需将编码后的水印信息进行分组. 由文献[8]分析结果可知,当嵌入载体图像的水印总能量一定时,用户数 K 的大小并不影响载体图像的 R_{PSN} 值,因此对水印信息进行分组时选择合适的用户数即可. 当水印系统满足一定不可见性,即在给定 R_{PSN} 值的

情况下,采用码长更长的扩频码对水印信息进行扩频,扩频后的水印信息能量增加,即嵌入载体图像的水印总能量增加,则嵌入强度 α 需要下降. 因此,实验综合考虑了可嵌入水印信息的载体图像 DCT 变换域中频系数的个数和载体图像的 R_{PSN} 值,选择合适的用户数 K 、扩频码长度 N 和嵌入强度 α . 实际应用中,一般认为载体图像的 R_{PSN} 值大于 38 dB 时,嵌有水印的载体图像在视觉上不易被察觉. 本实验嵌入水印信息的载体图像的 R_{PSN} 值为 40.9 dB,嵌入水印信息后的载体图像如图 1(c)所示. 用误码率(R_{BE})和相似性(NC)来衡量提取出的水印图像与原始水印图像的差别和相似度.



图1 载体图像和原始水印图像

Fig.1 Carrier image and original watermarking image

为验证所提算法的鲁棒性性能,本实验给出了基于 LDPC 码和 CDMA 扩频的水印方案,采用非规则 LDPC 码,码率为 1/4,校验矩阵的行重不固定,列重为 3,译码方案为 LLR-BP 译码. 该实验对嵌入水印的载体图像进行 JPEG 压缩攻击,图 2 为文献[6]方案、文献[15]算法、LDPC 编码方案和本文算法提取出的水印误码率结果. 从图 2 中可以看出,当压缩因子 $Q = 10, 20$ 的情况下,图像压缩率较大,本文算法的水印误码率均低于其他 3 种方法;而在压缩因子 $Q = 30, 40$ 时,此时 JPEG 压缩率减小. 文献[6]方案由于采用非盲水印检测的方法,即检测时需要原始载体图像,提取出的

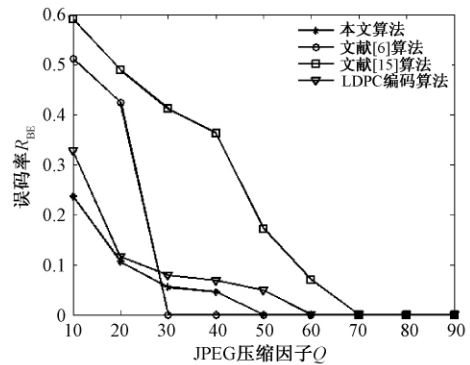


图2 不同算法在 JPEG 压缩攻击后提取出的水印误码率

Fig.2 BER during the JPEG compression attacks using different algorithms

水印误码率下降的较快;而本文算法采用盲检测的方式,此时检测出的正确水印比特不足以让 LT 码实现完全译码,因此误码率并没有降为零.实验结果表明,本文算法具有较好的抗 JPEG 压缩能力,且能有效抵抗高 JPEG 压缩.

表 1 给出了不同噪声和滤波攻击时 3 种方案提取出的水印图像和原始水印图像的相似值和误码率.实验中高斯低通滤波的窗宽为 4×4 ,标准差为 0.2,均值滤波和中值滤波的窗宽均为 3×3 .由表 1 实验结果可知,本文算法的抗高斯滤波和均值滤波性能低于文献 [15] 算法,这是因为在载体图像中嵌入水印信息相当于在图像中加入噪声,而本文算法水印信息仅嵌入了 DCT 变换域的某些中频系数.经过线性的均值滤波和高斯低通滤波后图像像素值改变较大导致相关检测时相关性降低,从而在水印提取时删除的水印比特过多,

而 LT 码不具有纠错能力,仅利用少量正确的水印信息不能完全译码,故其误码率增大.从表 1 可以看出,对水印化载体图像进行非线性的中值滤波时,与其他方案相比,本文算法具有较低的误码率.通过实验结果对比可知,嵌入水印的载体图像遭受不同程度噪声的攻击时,应用本文算法提取出的水印的误码率仍然很小,这表明该方案的抗噪声攻击的能力很强.

表 2 给出 3 种方案在剪切比分别为 65%、25%、50% 和中心剪切的剪切攻击下 3 种算法提取出的水印图像及误码率.实验中嵌有水印的载体图像被剪切的部分用白色像素替代.在剪切攻击时用白色像素填充的部分在水印相关检测时相关值低于判决阈值而被删除,提取时利用未被剪切部分检测出的正确水印信息进行 LT 码译码,恢复出原始水印信息.从表 2 可见,

















表 1 水印化图像在各种噪声和滤波攻击后提取出的水印相关值和误码率

Tab. 1 Correlation values and BER during all kinds of noise and filtering attacks on watermarked image

攻击方式及参数	本文算法		文献 15 算法		LDPC 码编码方案	
	NC	$R_{BE}/\%$	NC	$R_{BE}/\%$	NC	$R_{BE}/\%$
椒盐噪声(方差 $var=0.01$)	0.994 5	1.855 5	0.993 1	2.343 8	0.973 8	9.863 3
椒盐噪声(方差 $var=0.1$)	0.940 8	30.761 7	0.936 1	36.523 4	0.932 2	34.570 3
高斯噪声(方差 $var=0.05$)	0.961 1	16.308 6	0.944 3	27.636 7	0.953 0	21.191 4
高斯噪声(方差 $var=0.1$)	0.958 7	17.480 5	0.946 2	26.074 2	0.946 3	25.585 9
中值滤波(3×3)	0.948 9	22.656 3	0.932 2	57.515 9	0.9312	35.742 2
高斯低通滤波(4×4)	0.966 8	13.085 9	0.993 7	2.148 4	0.964 6	13.575 4
均值滤波(3×3)	0.915 0	43.554 7	0.972 7	10.058 6	0.925 4	38.769 5

表 2 不同剪切比例下提取出的水印图像及误码率

Tab. 2 The BER and recovered watermark image caused by crop attacks

剪切图像	本文算法		文献 [15] 算法		LDPC 码编码方案	
	水印	$R_{BE}/\%$	水印	$R_{BE}/\%$	水印	$R_{BE}/\%$
		10.742 2		19.628 9		30.078 1
		0		0.781 3		8.593 8
		0		7.714 8		0
		0		0.488 3		4.101 6

本文算法采用 LT 码编码方式优于 LDPC 码编码方案;与文献[15]算法的对比结果表明,本方案采用 CDMA 扩频方式具有更好的鲁棒性.表2实验结果表明,本文算法在剪切比低于 50% 时均可无差错地恢复出原始水印信息,具有很强的抗剪切攻击能力.

针对该方案的水印容量问题,实验时选择合适的扩频码长度 N 和嵌入强度 α ,保证载体图像嵌入不同大小的水印信息后仍然能满足一定的不可见性,嵌入的水印信息均为二值图像,实验中最大的水印图像大小为 64×64 .图3给出了该算法在不同水印嵌入量且没有攻击时提取出的水印信息的误码率.当水印嵌入量为 2^{12} bit 时,误码率为 0.041 5;水印嵌入量小于 2^{12} bit 时,均可无差错地提取出水印信息.实验结果表明,本文算法具有较大水印容量,且水印嵌入量较大时仍然具有较低的误码率,水印系统稳健性良好.

实验还通过考量 3 种水印方案在水印嵌入时所需的时间,以考察方案的复杂性和可实行性.3 种方案均嵌入 32×32 的二值图像水印信息,仿真 100 次得到水印嵌入时间的平均值.本文算法、文献[15]算法和 LDPC 码编码方案的水印嵌入时间分别为 72.7 s、65.3 s 和 344.2 s.本文算法采用 LT 码编码方式,较传统信道编码 LDPC 码的编码方案的水印嵌入时间大大减小,比文献[15]算法的水印嵌入时间略有增大.综合考虑可知,采用本文方案的水印系统具有较好的鲁棒性和较大的水印容量,说明本文方案具有一定的可行性.

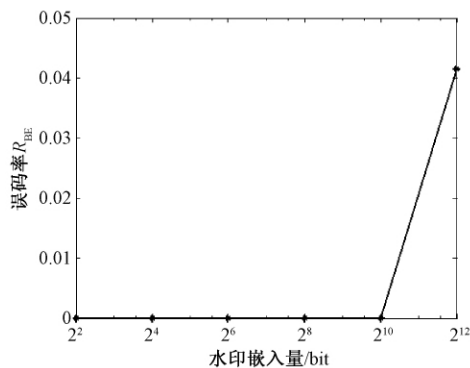


图3 嵌入量对比特错误率的影响

Fig.3 The affect of R_{BE} in different embedding quantity

5 结束语

笔者所提算法综合考虑了 CDMA 扩频水印模型各种指标之间的制约关系,并采用 LT 码对

水印信息进行预编码,与传统的信道编码技术相比具有更低的编译码复杂度;同时采用 CDMA 扩频方式大大提高了水印容量,水印提取时不需要原始载体图像和水印图像,从而实现了完全意义上的水印盲检测.本文算法中生成扩频码集和选择载体图像 DCT 变换域系数均需要密钥控制,因此,水印检测提取时需要原始密钥,增强了水印系统的安全性.实验结果表明,本文算法不仅复杂度低,具有较大的水印容量,而且在高 JPEG 压缩比、高剪切比和噪声攻击下有很好的鲁棒性,并对其他常见图像处理和攻击也有较好的鲁棒性.为了进一步提高算法的稳健性,今后可以采用性能更优异的其他度分布函数来提高 LT 码的编译码性能,降低译码的误码率.因 LT 码不具有纠错能力,未来也可以采用其他实用的喷泉码作为水印的预编码方案,以提高数字水印系统的鲁棒性.

参考文献:

- [1] COX I J, KILIAN J, LEIGHTON T. Secure spread spectrum watermarking for multimedia [J]. IEEE transactions on image processing, 1997, 6(12): 1673 - 1687.
- [2] KHALIL M, ADIB A. Embedding and extracting multiple watermarks in audio signals using CDMA [C]// 2014 9th International Conference on Intelligent Systems: Theories and Applications. Rabat: IEEE, 2014: 1 - 7.
- [3] JURGEN G, MUHAMMAD K S. CDMA based reversible and blind watermarking scheme for images [J]. International journal of future computer and communication, 2013, 2(2): 138 - 141.
- [4] KESHAV S R, DHEERENDRA S T. Digital watermarking schemes for authorization against copying or piracy of color images [J]. Indian journal of computer science and engineering, 2010, 1(4): 295 - 300.
- [5] SILVESTRE G C M, DOWLING W J. Embedding data in digital images using CDMA techniques [C]// Proceedings of the 7th IEEE International Conference on Image Processing. Vancouver, Canada: IEEE, 2000: 589 - 592.
- [6] 高琦,李人厚,王慧琴,等. 基于 Gold 码的扩频数字水印算法 [J]. 西安交通大学学报, 2004, 38(2): 68 - 75.
- [7] 许文丽,李磊,王育民. 抗噪声、几何失真和 JPEG 压缩攻击的鲁棒数字水印方案 [J]. 电子与信息学报, 2008, 30(4): 933 - 936.
- [8] 方艳梅,谷利民,黄继武. 利用边信息嵌入的 CDMA 水印信道性能研究 [J]. 电子学报, 2006, 34(1):

- 45 – 50.
- [9] OROZCO – VIVIAN L, SHAHRAM Y. Trapping sets of fountain codes [J]. IEEE communications letters, 2010, 14(8): 755 – 757.
- [10] KORUS P, BIALAS J, DZIECH A. A new approach to high-capacity annotation watermarking based on digital fountain codes [J]. Multimedia tools and applications, 2014, 68(1): 59 – 77.
- [11] GUO T C, LEI C, ZHENG H F. Analysis of the degree distribution for LT codes [J]. International journal of wireless and mobile computing, 2014, 7(3): 224 – 230.
- [12] CHEN C M, CHEN Y P, SHEN T C, et al. A practical optimization framework for the degree distribution in LT codes [J]. IEEE transactions on communications, 2013, 96, B(11): 2807 – 2815.
- [13] MICHAEL L. LT codes [C] // Proceedings of the 43rd Annual IEEE Symposium on the Foundations of Computer Science. Vancouver B C, Canada: IEEE, 2002: 271 – 280.
- [14] 克兢, 郭建新, 唐升. 一种基于 LT 码的抗剪切攻击信息隐藏算法 [C] // Proceedings of 2010 Asia-Pacific Conference on Information Network and Digital Content Security(2010APCID). 北京: 中国人工智能学会智能数字内容安全专业委员会, 2010: 4 – 5.
- [15] 张卫党, 李方圆, 陈美玲. 基于喷泉码的数字图像水印的研究 [J]. 郑州大学学报(工学版), 2014, 35(6): 52 – 55.
- [16] PARAH S A, ASHRAF S, ASHRAF A. Robustness analysis of a digital image watermarking technique for various frequency bands in DCT domain [C] // 2015 IEEE International Symposium on Nanoelectronic and Information Systems. Indore: IEEE, 2015: 57 – 62.

Blind Digital Watermarking Algorithm Based on Fountain Code and CDMA Spread Spectrum

ZHANG Weidang, WU Lin

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: In order to improve the robustness and watermark capacity of digital image watermarking system, this paper presented a digital image watermarking algorithm based on fountain code and CDMA spread spectrum techniques. LT code for the purpose of encoding the original watermark was adopted, and then CDMA spread spectrum techniques was used to generate the encoded watermark information. The encoded watermark information was embedded in the DCT domain of the carrier image. Blind recovery of the embedded data was achieved by using the character of cross-correlation function of the orthogonal spread spectrum code. The error watermark bits would be deleted by setting decision threshold and then using the correct watermark bits to decode, which could improve the robustness of digital watermarking system. The result showed the scheme proposed in this paper could embed larger quantities of information and raise robustness, especially it was robust against JPEG compression, noise and cropping attacks.

Key words: digital watermarking; CDMA; fountain code; DCT; robustness

基于混合差分演化的网络入侵检测算法

王耀光¹, 陈伟权¹, 吴镇邦¹, 秦勇², 黄翰³

(1. 广东省东莞市质量监督检测中心, 广东 东莞 523000; 2. 东莞理工学院 计算机学院, 广东 东莞 523000; 3. 华南理工大学 软件学院, 广东 广州 510006)

摘要: 基于机器学习方法的入侵检测算法是目前网络设备检测领域的研究热点. 网络入侵检测源数据的多样性是影响机器学习方法在该领域实际应用性能的主要因素. 研究通过设计多扰动向量混合差分演化算法, 稳定地优化了最小二乘支持向量机模型的关键参数; 在不增加测试集检测计算复杂性的前提下, 通过最优化参数的方式, 提高了最小二乘支持向量机算法入侵检测的精度和稳定性. KDD Cup 99 测试集的仿真实验结果显示, 所提出的基于混合差分演化的网络入侵检测算法比目前多种同类算法有着更好的平均性能.

关键词: 网络入侵检测; 测试稳定性; 混合差分演化; 最小二乘支持向量机

中图分类号: TP301.6 **文献标志码:** A **doi:** 10.13705/j.issn.1671-6833.2017.06.006

0 引言

入侵检测系统 IDS (intrusion detection system) 是计算机信息安全领域的一个重要分支. IDS 的特点是主动防御, 即对入侵行为进行预警, 关键技术是对入侵事件的识别和分类上. 近年来, 机器学习与智能计算相结合的方法成为了网络入侵检测研究的热门技术. Hu 等^[1] 利用粒子群算法对在线 Adaboost 的参数进行优化, 解决了动态分布式网络入侵问题. 刘羿^[2] 提出了蝙蝠算法优化神经网络, 提高了网络入侵检测的效率. 李振刚等^[3] 运用改进的蚁群算法来优化 SVM 参数, 提升了入侵检测的效率和准确率. 王亚等^[4] 研究发现通过优化 RBF 神经网络的参数, 可以有效降低特征维数和 RBF 神经网络输入节点数, 从而降低计算复杂度, 加快网络入侵检测速度. Dastanpour 等^[5] 系统研究了遗传算法在优化神经网络和支持向量机参数上的效果, 发现网络入侵检测源数据的多样性是影响演化算法实际应用性能的主要因素. 实际工程应用对网络入侵检测的准确率和响应速度要求较高, 因此, 高速、高准确率的要求是困扰智能计算方法应用于网络入侵检测的核心难题.

为了解决这一难题, 差分演化算法和网络入侵检测的结合成为了研究热点. 马琰等^[6] 将混沌差分算法用于网络入侵检测, 降低了检测的误差; 边根庆等^[7] 将免疫克隆与差分进化相统一, 为进化算法在网络入侵检测中的实际应用做了理论铺垫; Sailaja 等^[8] 基于差分演化算法对网络入侵检测问题进行了研究. 考虑到网络入侵检测源数据的多样性, 笔者采用了多扰动向量的混合差分演化算法, 对最小支持向量机 (LSSVM) 的大量关键参数进行优化, 通过对于差分演化算法的改进, 使得优化后的 LSSVM 提高了网络入侵检测的准确率和稳定性.

1 基于最优参数 LSSVM 网络入侵检测的可行性分析

由于入侵事件种类的多样性和稳定性, 网络入侵检测可以建模为一个模式分类问题. 许多网络入侵检测技术都应用最小支持向量机 (LSSVM) 来进行入侵事件的分类; 然而, LSSVM 的分类准确率受其多个参数的影响. 通过参数优化来提高网络入侵检测精度的研究是目前学术界研究的热点, 其技术路线如图 1 所示.

LSSVM 算法定义如公式 (1):

收稿日期: 2017-05-20; 修订日期: 2017-07-18

基金项目: 国家自然科学基金资助项目 (61370102); 广东省高等院校学科与专业建设专项资金建设项目 (2013KJCX0178)

作者简介: 王耀光 (1964—), 男, 广东东莞人, 广东省东莞市质量监督检测中心高级工程师, 主要从事信息技术设备安全质量评估方面的研究, E-mail: wyg@gddqt.com.

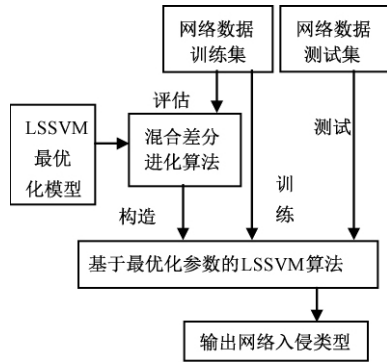


图1 基于LSSVM模式分类的网络入侵检测技术路线图

Fig. 1 Technology roadmap of network intrusion detection based on LSSVM model classification

$$\begin{cases} \min J(\omega, \xi) = \frac{1}{2}(\|\omega\|^2 + C \sum_{i=1}^n \xi_i^2); \\ \text{s. t. } y_i [\omega^T \varphi(x_i) + b] - 1 + \xi_i = 0. \end{cases} \quad (1)$$

可以通过引入拉格朗日乘子将式(1)的模型变化为式(2)的实优化问题:

$$\min L = J(\omega, \xi) - \sum_{i=1}^n a_i \{y_i \cdot [\omega^T \varphi(x_i) + b] - 1 + \xi_i\}. \quad (2)$$

一般的做法是将式(2)对 ω 、 ξ_i 、 b 和 a_i 求偏导,通过消去 ω 和 ξ_i 得到关于 b 和 a_i 的方程组,用数值计算的方法可以求解方程组并得到最优化的 b 和 a_i , $i=1,2,\dots,n$, n 为采样的规模.因为数值计算的方法所需计算时间长且复杂性较大,一些LSSVM应用为了节省计算时间而采用了近似算法,如粒子群优化算法等.但是,因为网络入侵检测的实时性要求较高,故而采用更高效的连续优化启发式算法来求得最优的 b 和 a_i .

在实际工程应用中可以运用训练集样本来评估LSSVM优化的效果.如果用于评估的样本足够描述网络入侵数据的特征,那么优化后的LSSVM模型可以高精度地分类出网络入侵事件.由于最优参数的LSSVM算法在实际检测过程中不再需要增加额外的计算量,因此,图1所示的方法在计算时间复杂性上是可行的.

2 改进算法设计

考虑到网络入侵检测的实时性与精确性要求,因此采用了差分演化算法(DE)^[9]作为优化LSSVM参数的核心技术,而DE是目前解决连续优化问题最有效的算法之一.虽然DE算法在单目标连续优化问题的求解上表现出了较强的性能,但是在求解式(2)的优化问题时容易陷入局

部最优解和收敛慢的困境,而扰动向量是对DE算法性能影响最大的一个因素.因此,笔者通过采用多个扰动向量设计扰动向量池来实现混合差分变异规则,从而提高算法的全局搜索能力和收敛速度.

2.1 基于扩展扰动向量池的混合差分演化算法

差分演化算法的目标是给不断演化的 N 个 D 维的向量找到全局最优,式(2)的LSSVM优化参数 b 和 a_i 可编码为 $X_i^G = \{x_{i,1}^G, x_{i,2}^G, \dots, x_{i,D}^G\}$, $i=1,2,\dots,N$.其中, G 为当前种群迭代次数; D 为目标函数维度大小.初始的向量分布最好均匀地分布在整个搜索空间内,搜索空间的最大和最小边界被预先设定: $X_{\min} = \{x_{\min,1}, x_{\min,2}, \dots, x_{\min,D}\}$ 和 $X_{\max} = \{x_{\max,1}, x_{\max,2}, \dots, x_{\max,D}\}$.借助文献[4]的方法可以将网络数据的字符映射为数值, X_{\min} 和 X_{\max} 根据这些数值的范围确定.

考虑到网络数据的多样性与入侵事件类型的稳定性,笔者设计了混合策略用来改进差分演化算法的性能.在混合策略中,将针对每一个混合步骤个体的变异设计多个扰动向量来生成新个体进入下一代,通过混合策略可以计算多个扰动向量的新子代,然后选取最优的进入下一代.混合策略可以被分为两类:①计算每一个扰动向量产生子代的最优值,如果其中有优于父代的则选取最优的一个进入下一代,反之继续沿用父代个体;②构建一个选择模型用来预测每一代应该通过哪一个扰动向量来产生新个体.通常一个好的预测模型可以在大多数迭代过程中选择最优的扰动向量.候选扰动向量的列表如式(3)~(15)所示.

$$V_i^G = X_{\text{best}}^G + F(X_{r1}^G - X_{r2}^G); \quad (3)$$

$$V_i^G = X_{\text{best}}^G + F(X_{r1}^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G); \quad (4)$$

$$V_i^G = X_{r1}^G + F \cdot (X_{r2}^G - X_{r3}^G); \quad (5)$$

$$V_i^G = X_{r1}^G + F(X_{r2}^G - X_{r3}^G) + F(X_{r4}^G - X_{r5}^G); \quad (6)$$

$$V_i^G = X_i^G + F(X_{\text{best}}^G - X_i^G) + F(X_{r1}^G - X_{r2}^G); \quad (7)$$

$$V_i^G = X_{r1}^G + F(X_{\text{best}}^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G), \quad (8)$$

式中: $F(X_{r2}^G - X_{r3}^G)$ 是一个权重的差分向量算子; X_{best}^G 是第 G 代的最优个体; F 是一个大于零的控制向量.式(3)~(8)的扰动向量设计参考了Storn和Price提出的差分演化算法扰动向量家族^[9].

由于公式(2)的优化模型由网络数据的样本向量决定,基于网络数据的多样性,图1所示的技术路线将造成优化模型的动态变化,因此,单一的优化策略难以有效地解决该问题.与经典LSSVM

参数优化算法不同,混合使用多种扰动向量可以增加差分演化算法种群的多样性.因此,可以参考基因算法的情况,每个个体可以选取同样的个体来产生新的子代,这在数学上增加了个体的多样性,参见图2.

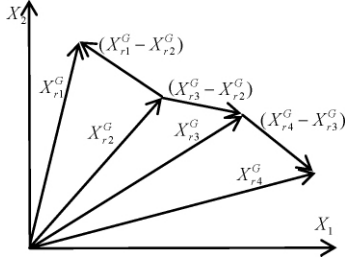


图2 在2维空间内解释DE扰动向量组件

Fig.2 Explained disturbance vector components of DE in the two-dimensional space

新的扰动向量生成策略增加了种群多样性的潜力, $(X_{r3}^G - X_{r2}^G)$ 就是增加的差分向量的选项.针对网络样本数据的多样性,笔者扩展设计了新的扰动向量,如式(9)至(15)所示:

$$V_i^G = X_{best}^G + F(X_{r1}^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G) + F(X_{r5}^G - X_{r6}^G); \quad (9)$$

$$V_i^G = X_{r1}^G + F(X_{r2}^G - X_{r3}^G) + F(X_{r4}^G - X_{r5}^G) + F(X_{r6}^G - X_{r7}^G); \quad (10)$$

$$V_i^G = X_{best}^G + F(X_i^G - X_{r1}^G) + F(X_{r2}^G - X_{r3}^G); \quad (11)$$

$$V_i^G = X_i^G + F(X_{r1}^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G); \quad (12)$$

$$V_i^G = X_{r1}^G + F(X_i^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G); \quad (13)$$

$$V_i^G = X_{r1}^G + F(X_{best}^G - X_{r1}^G) + F(X_{r2}^G - X_{r3}^G) + F(X_{r4}^G - X_{r5}^G); \quad (14)$$

$$V_i^G = X_{r1}^G + F(X_{best}^G - X_{r2}^G) + F(X_{r3}^G - X_{r4}^G) + F(X_{r5}^G - X_{r6}^G). \quad (15)$$

式(9)~(15)提出的扰动向量兼顾了局部搜索以及全局搜索的能力,在多样性的优化问题上比较容易取得稳定的效果,可以弥补 Storn 和 Price 提出的差分演化扰动向量家族的不足.

2.2 改进差分演化算法的流程

按照式(3)~(15)选取扰动向量,流程如图3所示.启发式步骤包括:在差分演化算法初始化时,随机从扰动向量池内选择一个扰动向量;运行差分演化算法时,如果所选的扰动向量没有进一步优化公式(2)的问题,重新随机选择一个不同于之前所选的扰动向量.

当基于新网络数据的优化问题生成时,即某个扰动向量不能使差分演化算法得到更优解时,可以重新选择扰动向量池内的扰动向量.而扰动向量可以根据不同的优化模型来选取,式(3)~

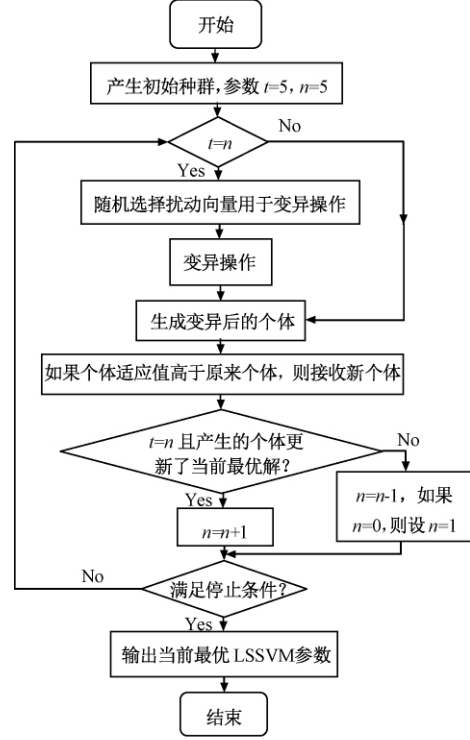


图3 改进差分演化算法的流程

Fig.3 The process of improved differential evolution algorithm

(15) 提供了备选扰动向量更新公式,它们都具有良好的可装卸性.通过拆卸和组装混合扰动向量池的扰动向量组合,可以更有效方便地解决不同类型网络数据在图1技术路线下形成的优化问题:如单峰问题适合使用 best 驱动的扰动向量;多峰问题适合利用 rand 偏移来跳出局部最优解.

3 仿真实验

3.1 实验设置

本实验采用 KDD CUP 99 数据进行离线测试,验证基于混合差分演化的网络入侵检测算法(简记为 HDE-SVM)的效率.笔者提出的 HDE-LSSVM 算法包含了差分演化和最小支持向量机两部分,参数设置如下:LSSVM 公式(2)中的参数由 HDE 算法优化确定,采样规模 n 设置为 30.根据文献[9-10]建议,HDE 种群规模 N 设置为 100,惯性权重 $w = 0.5$, $CR = 0.9$.

3.2 实验结果与分析

笔者提出的 HDE-SVM 算法采用 6 个节点进行测试,通过 KDD CUP 99 的数据测试对比了 PSO-SVM、PSO-LSSVM、投票方法^[11]、DE-LSSVM、SaDE-LSSVM^[12] 和 HDE-LSSVM 的性能,研究结果以检测准确率和误报率作为对比指标,实验结果如表1所示.

表 1 PSO-SVM、PSO-LSSVM、投票方法、DE-LSSVM、SaDE-LSSVM 和 HDE-LSSVM 的对比实验结果

Tab. 1 The contrast experiment results of PSO-SVM, PSO-LSSVM, voting method, DE-LSSVM, SaDE-LSSVM and HDE-LSSVM												
节点	PSO-SVM		PSO-LSSVM		投票方法		DE-LSSVM		SaDE-LSSVM		HDE-LSSVM	
	准确率	误报率	准确率	误报率	准确率	误报率	准确率	误报率	准确率	误报率	准确率	误报率
节点 1	99.59	0.38	99.76	0.44	99.96	7.95	98.45	0.51	99.13	0.48	99.92	0.37
节点 2	99.70	0.44	99.78	0.44	99.96	7.95	98.83	0.55	99.21	0.46	99.93	0.39
节点 3	99.67	0.43	97.13	0.42	99.96	7.95	98.56	0.58	99.25	0.45	99.95	0.40
节点 4	99.78	0.44	99.78	0.44	99.96	7.95	98.36	0.52	99.18	0.49	99.85	0.40
节点 5	99.65	0.41	99.78	0.45	99.96	7.95	98.78	0.55	99.24	0.44	99.34	0.39
节点 6	99.74	0.48	97.13	0.42	99.96	7.95	97.89	0.53	99.31	0.47	99.92	0.41
平均	99.69	0.43	98.89	0.44	99.96	7.95	98.48	0.54	99.22	0.47	99.82	0.39

表 1 显示, HDE-LSSVM 在 6 个节点的网络测试数据上有着比较稳定的检测准确率和误报率。虽然, HDE-LSSVM 并没有在每个节点相对其他方法取得最高准确率, 但是总体的平均准确率次于投票方法。稳定的高准确率说明, HDE 求得的最优 LSSVM 模型可以达到稳定且准确分类的水平, 提高了 LSSVM 的鲁棒性。值得一提的是, HDE-LSSVM 的误报率相对较低, 这说明了 HDE 比 PSO 达到了更高精度的优化效果。

除此以外, 还进行了标准差分演化算法(DE) 和自适应差分演化算法(SaDE) 优化 LSSVM 的仿真实验。实验结果表明, DE-LSSVM 和 SaDE-LSSVM 在检测准确率和误报率上都劣于 PSO-LSSVM 方法, 这一现象也说明了 Dastanpour 等的观点: 网络入侵检测源数据的多样性影响了演化算法优化 SVM 的精度。笔者提出的 HDE-LSSVM 算法则弥补了这一缺陷, 显著提高了网络入侵检测的效率。

4 结论

笔者从优化最小二乘支持向量机的关键参数入手, 用差分演化算法取得了更高、更稳定的优化效率; 并针对网络数据的多样性, 研究设计了多种扰动向量丰富了差分演化算法的扰动向量池, 实现了 LSSVM 的自适应参数调优。实验结果表明, 多扰动向量的策略大大提高了优化性能, 并且使得 LSSVM 在网络入侵检测上有更稳定的平均性能。

参考文献:

- [1] HU W M, GAO J, WANG Y G, et al. Online adaboost-based parameterized methods for dynamic distributed network intrusion detection [J]. IEEE transactions on cybernetics, 2014, 44(1) : 66 - 82.
- [2] 刘羿. 蝙蝠算法优化神经网络的网络入侵检测 [J]. 计算机仿真, 2015, 32(2) : 311 - 314.
- [3] 李振刚, 甘泉. 改进蚁群算法优化 SVM 参数的网络入侵检测模型研究 [J]. 重庆邮电大学学报(自然科学版), 2014, 26(6) : 785 - 789.
- [4] 王亚, 熊焰, 龚旭东, 等. 基于混沌 PSO 算法优化 RBF 网络入侵检测模型 [J]. 计算机工程与应用, 2013, 49(10) : 84 - 87.
- [5] DASTANPOUR A, IBRAHIM S, MASHINCHI R, et al. Comparison of genetic algorithm optimization on artificial neural network and support vector machine in intrusion detection system [J]. Open systems, 2014, 77(10) : 72 - 77.
- [6] 马琰, 闫兵. 基于混沌差分优化算法的网络入侵检测系统 [J]. 科学技术与工程, 2013, 13(36) : 10967 - 10970.
- [7] 边根庆, 赵宏, 张维琪, 等. 基于免疫克隆与差分进化的入侵检测方法 [J]. 微电子学与计算机, 2012, 29(5) : 124 - 128.
- [8] SAILAGA M, KUMAR R K, MURTY P S R. Intrusion detection model based on differential evolution [J]. International journal of computer applications, 2011, 36(6) : 10 - 13.
- [9] STORN R, PRICE K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces [J]. Journal of global optimization, 1997, 11(4) : 341 - 359.
- [10] BREST J, ŽUMER V, MAUCEC M. Self-adaptive differential evolution algorithm in constrained real-parameter optimization [J]. IEEE congress on evolutionary computation, 2006, 98: 215 - 222.
- [11] DENG W, YANG X, ZOU L, et al. An improved self-adaptive differential evolution algorithm and application [J]. Chemometrics and intelligent laboratory systems, 2013, 128(15) : 66 - 76.
- [12] YU C C, CHEN J, HUANG Q, et al. A new hybrid differential evolution algorithm with simulated annealing and adaptive Gaussian immune [C] // 2012 8th IEEE International Conference on Natural Computation (IC-NC). Chongqing, China: IEEE, 2012: 600 - 607.

(下转第 49 页)

精英反向学习的单纯形交叉布谷鸟搜索算法

林要华, 王李进

(福建农林大学 计算机与信息学院, 福建 福州 350002)

摘要: 提出一种加强搜索能力的改进布谷鸟搜索算法, 该算法采用精英反向学习策略促使 Lévy Flights 随机走动中的部分精英个体进行反向搜索, 以避免搜索新个体的趋同性; 并采用单纯形交叉操作在 Biased 随机走动中随机选择一个个体进行精细搜索, 以降低搜索的盲目性以及低效性. 另外, 提出的算法采用混沌映射模型实现发现概率参数的自适应控制. 仿真实验结果表明, 该算法能够总体上有效改善算法的搜索能力和收敛速度.

关键词: 布谷鸟搜索算法; 单纯形交叉; 反向学习; 混沌映射

中图分类号: TP301

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.033

0 引言

近年来, 涌现诸如人工蜂群算法^[1]、萤火虫算法^[2]、群搜索算法^[3]、布谷鸟搜索算法^[4-5]等一些新颖算法. 布谷鸟搜索算法(cuckoo search, CS)在求解函数优化问题上能获得较好的性能, 同时具有运算简单和参数较少的特点, 这些使得该算法已成为研究的热点. 一方面, 一些学者针对算法中的随机走动进行了研究与改进, 文献[6-8]开展了 Lévy Flights 随机走动及其步长的相关工作; 文献[9-10]则对加强 Biased 随机走动的搜索方式进行了研究. 另一方面, 一些学者通过混合其他优化算法以提高算法的搜索能力^[11-13].

上述研究容易导致种群的同质性问题, 而且种群中的当前最优个体无法从自身获得任何步长信息, 这也限制了该个体的搜索能力. 另外, Biased 随机走动中的随机交叉搜索方式可能会导致盲目或者低效搜索问题. 因此, 笔者针对上述两个问题, 引进解对称和单纯形交叉策略, 以改善 CS 算法的收敛速度以及求精能力, 并提出了精英反向学习的单纯形交叉布谷鸟搜索算法(elite opposition-based learning based simplex crossover cuckoo

search algorithm, ESCS).

1 布谷鸟搜索算法

布谷鸟搜索算法是 Yang 和 Deb^[5-6]基于布谷鸟的繁殖策略而构建的算法. 该算法主要包含 3 种部件: Lévy Flights 随机走动、Biased 随机走动以及贪婪选择. 对于求解连续函数优化问题, CS 算法首先在问题的解空间 $[x_{j,\min}, x_{j,\max}]$ ($j=1, 2, \dots, D$) 内随机初始化种群.

$x_{i,j,0} = x_{i,j,\min} + r(x_{i,j,\max} - x_{i,j,\min})$, $i=1, 2, \dots, N$, (1)
式中: $r \in [0, 1]$ 是缩放因子; N 是种群大小.

在初始化之后, CS 算法进入进化阶段. 在进化过程中, Lévy flights 随机走动和 Biased 随机走动被分别用于搜索新的个体. 在每个随机走动之后, CS 算法比较每个当前个体和新搜索个体的适应值, 贪婪选择适应值较好的个体进入种群.

在 Lévy flights 随机走动中, 种群以当前获得的最优个体为导向进行进化, 其形式可表达为

$$X_{i,G+1} = X_{i,G} + \alpha_0 \frac{\varphi \cdot u}{|v|^{1/\beta}} (X_{i,G} - X_{\text{best}}). \quad (2)$$

式中: $X_{i,G}$ 、 $X_{i,G+1}$ 和 X_{best} 分别表示第 G 代的第 i 个当前个体、新个体以及当前最优个体; α_0 为缩放因子(取 0.01); μ, v 服从正态分布; $\beta=1.5$; 以及

收稿日期: 2017-02-22; 修订日期: 2017-07-29

基金项目: 福建省自然科学基金资助项目(2016J01280); 福建省教育厅资助项目(JB09114)

作者简介: 林要华(1979—), 男, 福建宁德人, 福建农林大学讲师, 主要从事智能算法及其应用研究.

通信作者: 王李进(1977—), 男, 福建泉州人, 福建农林大学副教授, 博士, 主要从事智能算法、智能信息处理研究,

E-mail: lijn wang@fafu.edu.cn.

$$\varphi = \left(\frac{\Gamma(1 + \beta) \cdot \sin\left(\frac{\pi \times \beta}{2}\right)}{\Gamma\left(\frac{1 + \beta}{2}\right) \cdot \beta \cdot 2^{\frac{\beta-1}{2}}} \right)^{1/\beta}, \quad (3)$$

其中, Γ 表示伽马(Gama) 方程.

在 Biased 随机走动中, 为了搜索新的个体, 当前个体随机选择两个个体, 各维按一定的概率接受它们的步长信息, 具体如下:

$$x_{i,j,G+1} = \begin{cases} x_{i,j,G} + r(x_{m,j,G} - x_{n,j,G}), & \text{if } \text{rand}(0,1) > p_a; \\ x_{i,j,G}, & \text{otherwise,} \end{cases} \quad (4)$$

式中, $x_{m,j,G}$ 、 $x_{n,j,G}$ 表示第 G 代的第 m 和第 n 个解; r 是缩放因子; $p_a \in [0,1]$, 是发现概率.

2 提出的改进算法

针对 CS 算法的搜索新个体的趋同性以及搜索低效率问题, 提出一种基于精英反向学习的单纯形交叉布谷鸟搜索算法, 这种算法采用精英反向学习、单纯形交叉以及发现概率 p_a 的自适应控制策略.

2.1 精英反向学习

基于群体搜索的智能算法的迭代过程是一个搜索区域变换到另一个搜索区域的过程^[14]. 反向学习(opposition-based learning, OBL)^[15]是实现区域变换搜索的一种模式. 文献[14]概括了4种常见反向学习模型.

$$x^* = k(a + b) - x, \quad (5)$$

式中: a 、 b 为搜索空间的上下界; k 为实数.

当 $k = 0$ 时, 得解对称的一般反向学习模型, 即

$$x^* = -x. \quad (6)$$

当 $k = 1/2$ 时, 得区间对称的一般反向学习模型, 即

$$x^* = \left(\frac{a + b}{2} \right) - x. \quad (7)$$

当 $k = 1$ 时, 得反向学习模型, 即

$$x^* = (a + b) - x. \quad (8)$$

当 $k = r$ 时, r 为 $[0,1]$ 之间的随机数时, 得到随机的一般反向学习模型, 即

$$x^* = r(a + b) - x. \quad (9)$$

2.2 单纯形交叉

单纯形交叉(Simplex Crossover, SPX)^[16]是一种基于交叉操作的局部搜索技术, 常用于加强个体的求精, 并呈现出潜在的性能^[17].

算法1 SPX 的流程如下.

步骤1: 计算 n 个父个体 $x_i (i = 1, \dots, n)$ 的中心位置 O .

$$O = \frac{1}{n} \sum_{i=1}^n x_i.$$

步骤2: 生成 $n-1$ 个随机数 r_i , $r_i = u^{\frac{1}{i+1}}$, $i = 1, \dots, n-1$; u 为 $[0,1]$ 之间的随机数.

步骤3: 计算 $y_i = O + \varepsilon(x_i - O)$, $i = 1, \dots, n$, ε 为扩张因子, $\varepsilon = 1.0$.

$$C_i = \begin{cases} 0, & i = 1; \\ r_{i-1}(y_{i-1} - y_i + C_{i-1}), & i = 2, \dots, n \end{cases}$$

步骤4: 生成子个体 $C = y_n + C_n$.

2.3 发现概率 p_a 的自适应控制

式(4)中的 p_a 控制着个体获取外部信息的概率. 越小的发现概率使得个体接受外部信息的机会越大, 这有利于搜索的全局性; 反之, 有利于个体的求精能力. 显然, 动态控制 p_a 有利于改善搜索部件的效率. 因此, 与文献[14]类似, ESCS 采用高斯混沌映射模型控制 p_a 参数, 见式(10), 其中初始 $p_a(0) = 0.7$.

$$p_a(i+1) = \begin{cases} 1, & p_a(i) = 0; \\ 1/\text{mod}(p_a(i), 1), & p_a(i) \neq 0 \end{cases} \quad (10)$$

2.4 ESCS 算法框架

算法2 给出了 ESCS 算法的流程. 需要指出的是在步骤2.1中, 除了采用式(8)外, 也可采用式(6)、式(7)和式(9)的反向学习模型搜索新的个体, 并引入参数 M 控制精英个体规模. 本文简单采用种群中的前 M 个最优个体作为精英个体. 另外, ESCS 选择待交叉个体以及当前种群中的 $K-1$ 个随机个体, 组合成 K 个个体作为单纯形交叉的父个体.

算法2 ESCS 的流程如下.

步骤1: 采用公式(1)初始化种群 $X_{ij} (i = 1, \dots, N; j = 1, \dots, D)$;

步骤2: 当迭代终止条件未满足;

步骤2.1: 的伪代码如下.

①For $i = 1$ to N

②if X_i 是精英个体而且精英个数 $\leq M$

③采用式(8)反向学习搜索个体 V_i

④Else

⑤采用式(2)搜索个体 V_i

⑥Endif

⑦如果 V_i 的适应值优于 X_i , 则用 V_i 替换 X_i

⑧Endfor

步骤2.2: 伪代码如下.

①从 $[1, N]$ 中随机产生待用单纯形交叉个体 indSPX

②选择 n 个体成为父个体

③根据式(10)获得 p_a

- ④For $i = 1$ to N
 - ⑤if $i = \text{indSPX}$
 - ⑥采用算法1生成候选子个体 V_i
 - ⑦else
 - ⑧采用公式(4)生成候选子个体 V_i
 - ⑨Endif
 - ⑩如果 V_i 的适应值优于 X_i , 则用 V_i 替换 X_i
 - ⑪Endfor
- 步骤2.3: 更新当前最优个体;
- 步骤3: 输出最优个体.

3 仿真与结果

为了分析 ESCS 算法的性能, 选择文献 [18] 中的 10 个旋转变换函数作为算法的测试函数. 其中, $F_1 \sim F_5$ 为单峰函数; $F_6 \sim F_{10}$ 为多峰函数. 更详细的测试函数内容见文献 [18]. 同时, 为了公平评价算法的性能, 笔者采用平均适应值误差及其收敛曲线图评价准则. 其中, 适应值误差 E 计算见式 (11).

$$E = f(X) - f_{\min}, \quad (11)$$

式中: X 为算法得到的解, 是已知的最优适应值. 从公式可以看出, 适应值误差越小则性能越优.

3.1 参数分析

ESCS 算法引入参数 M 和 K 分别用于控制参与精英反向学习个体的规模和参与单纯形交叉的父个体规模. 由于 M 和 K 存在不同的取值范围, 这就形成规模较大的组合, 因此, 本实验中假设其中一个参数取常量情况下分析另一个参数.

首先, 为了分析参数 M 对算法性能的影响, 图1给出了 ESCS 算法在 K 为常量以及不同 M 时求解 10 个测试函数 ($D = 30$) Friedman 的检验平均排序. 其中, 每个 M 取值所对应的算法在每个测试函数上独立运行 25 次, 种群大小 $N = 30$, 最大函数评价次数为 $10\,000 \times D$.

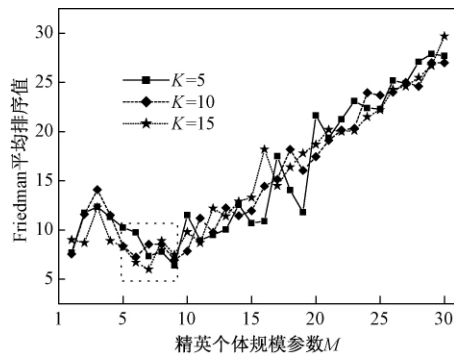


图1 不同 M 时的平均排序值

Fig.1 Mean rank under different M

从图1可以看出, M 取较小时, 算法总体上获得较好性能. 这是因为大部分的个体仍采用式 (2) 搜索新个体时, 种群的小部分精英个体参与反向学习确保了新个体的趋同性, 从而加强其搜索能力. 另外, 从图1可知, 当 $M \in [6, 9]$ 时, 算法的性能总体最优; 综合不同 K 值情况下的性能可知, $M = 9$ 较为理想.

图2给出当 $M \in [6, 9]$ 时, 取不同 K 值的算法求解函数的 Friedman 检验的平均排序. 从图2中可以看出, M 不同时, 算法获得最优的平均排序的 K 值是不同的. 但是, 从图2中的水平线 (平均排序值 = 13.3) 可知, 只有 $K = 15$ 所对应的 4 个 M 的平均排序都处于水平线下, 而 K 取其他值所对应 M 的平均排序少数低于水平线下, 这说明了 $K = 15$ 是较为理想的.

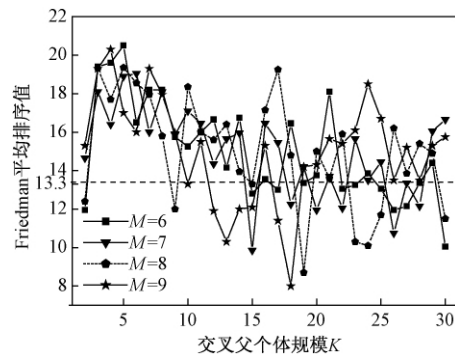


图2 不同 K 时的平均排序值

Fig.2 Mean rank under different K

3.2 性能分析

表1给出了 ESCS 算法和 CS 算法在 $D = 30$ 维上的 “ $E_{\text{avg}} \pm E_{\text{sd}}$ ” (E_{avg} 为平均适应值误差, E_{sd} 为标准方差). 其中, 每个算法在每个测试函数上独立运行 25 次, 两种算法的种群大小 $N = 30$, 而 ESCS 算法的 $M = 9, K = 15$, 最大函数评价次数为 $10\,000 \times D$. 另外, 表1也给出了两种算法在 0.05 显著水平下的 Wilcoxon 符号秩检验, “+”表示 ESCS 算法显

表1 CS 和 ESCS 求解 $D = 30$ 维的性能

Tab.1 Performance of CS and ESCS for solving $D = 30$

函数	CS		差异 性	ESCS	
	$E_{\text{avg}} \pm E_{\text{sd}}$			$E_{\text{avg}} \pm E_{\text{sd}}$	
F_1	2.52E-30	$\pm 7.51\text{E}-30$	=	2.02E-30	$\pm 1.01\text{E}-29$
F_2	7.38E-03	$\pm 8.82\text{E}-03$	+	2.08E-09	$\pm 4.88\text{E}-09$
F_3	2.22E+06	$\pm 6.19\text{E}+05$	+	1.27E+05	$\pm 1.06\text{E}+05$
F_4	1.38E+03	$\pm 7.33\text{E}+02$	+	1.88E+01	$\pm 1.91\text{E}+01$
F_5	3.09E+03	$\pm 6.75\text{E}+02$	+	5.83E+02	$\pm 4.32\text{E}+02$
F_6	2.22E+01	$\pm 2.40\text{E}+01$	+	4.78E-01	$\pm 1.32\text{E}+00$
F_7	7.09E-04	$\pm 1.62\text{E}-03$	-	1.66E-02	$\pm 1.32\text{E}-02$
F_8	2.09E+01	$\pm 4.34\text{E}-02$	=	2.09E+01	$\pm 1.04\text{E}-01$
F_9	2.74E+01	$\pm 5.44\text{E}+00$	+	2.05E+01	$\pm 5.97\text{E}+00$
F_{10}	1.62E+02	$\pm 2.59\text{E}+01$	+	5.81E+01	$\pm 1.22\text{E}+01$

著优于 CS 算法,“-”表示 CS 算法显著优于 ESCS 算法,而“=”表示两算法不存在显著差异。

从表 1 可以看出,ESCS 算法总体上可以改善求精能力.由 Wilcoxon 统计检验的结果分析可知,虽然在 F_7 测试函数上,ESCS 算法的求解能力未显著优于 CS 算法,但在其他测试函数上,ESCS 算法性能显著优于 CS 算法。

为了进一步比较算法的收敛性能,图 3 形化

展示 CS 和 ESCS 算法的收敛过程,其中,收敛曲线图中的平均适应值误差均是变换后的平均适应值误差,变换函数是 \lg . 从图 3 可以看出,对于 ESCS 算法获得的适应值优于 CS 算法的函数,ESCS 算法同时具有较快的收敛速度,在迭代后期具有较好的求精能力,可见图 3 中的(a)~(f)、(h)、(i)和(j).对于 F_7 函数的收敛过程见图 3(g),ESCS 算法在前期具有较快的收敛,但后期陷入局部最优。

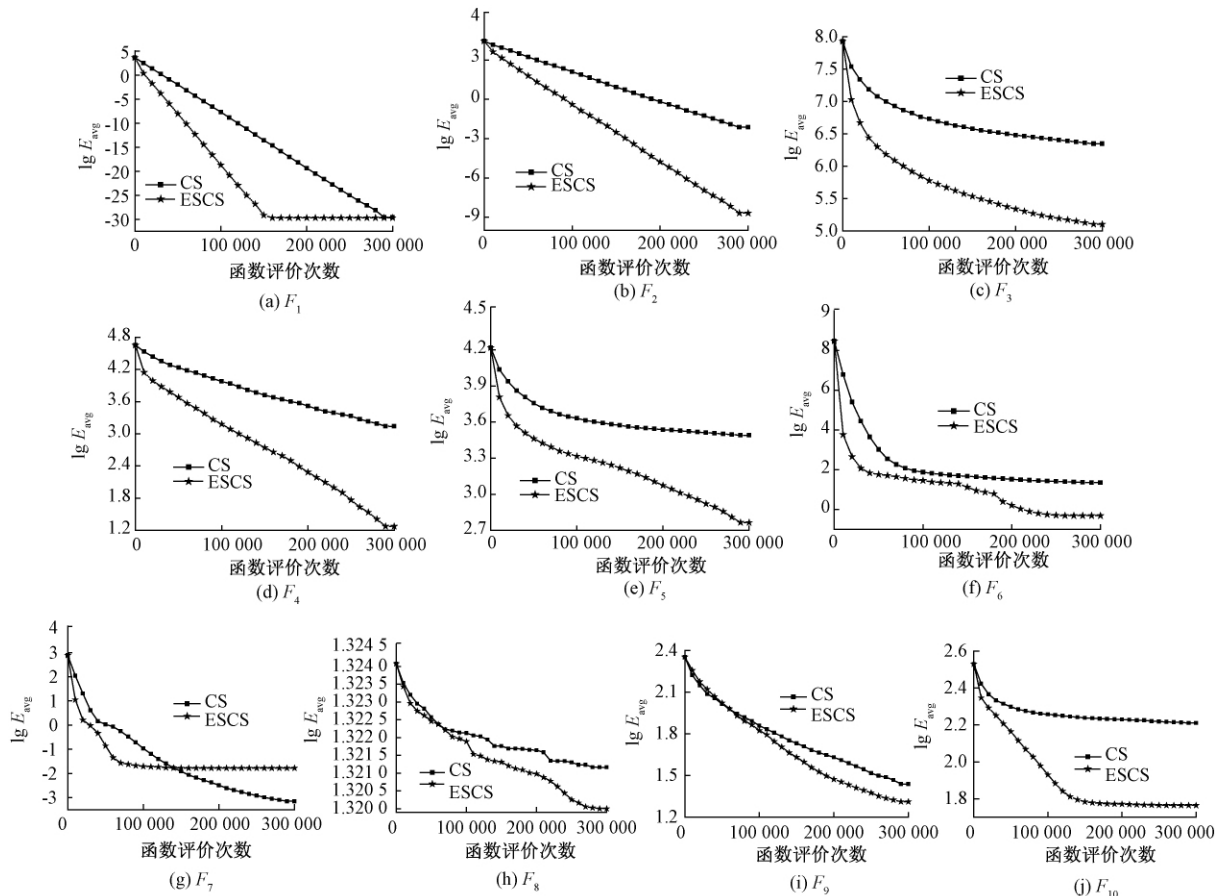


图 3 CS 和 ESCS 的收敛曲线

Fig. 3 Convergence curves of ESCS and CS

综合上述的实验结果,相对于 CS 算法,ESCS 算法对于大多数函数具有较强的求精能力以及较快的收敛速度。

3.3 ESCS 与其他改进的 CS 算法比较

为了进一步观察 ESCS 与其他改进的 CS 算法的竞争优势,表 2 列出 ESCS 算法与 CSPSO^[11]、OLCS^[12]以及 CCS^[8]在 $D=30$ 维空间上的求精能力比较结果,其中用于对比算法的实验数据来源于文献[8]。表中加粗数据表示相应算法在该函数获得最优性能,可以得出,不同算法在求解给定的测试函数上具有不同的优势;ESCS 算法在 F_4 、 F_5 、 F_6 和 F_{10} 函数上获得最好的解质量。另外, CSPSO、

OLCS、CCS 和 ESCS 4 种算法的 Friedman 检验结果分别为 3.2、3.3、1.8、1.7。从检验结果可知,ESCS 具有最小的平均排序值,这说明 ESCS 算法总体性能最优,其次依次为 CCS、CSPSO 和 OLCS。

3.4 不同反向学习模型对算法性能影响分析

为了分析不同反向学习模型对算法性能的影响,表 3 给出了 ESCS 在不同反向学习模型下求解测试函数所获得的平均误差,而表 4 列出相关算法的 Friedman 检验结果。其中,ESCS₆、ESCS₇ 和 ESCS₉ 表示 ESCS 算法分别采用式(6)、式(7)以及式(9)的反向学习模型。

从表 3 可知,不同反向学习模型使算法求

表2 ESCS与CSPSO、OLCS以及CCS的性能比较

Tab.2 Performance comparison of ESCS with CSPSO, OLCS and CCS

函数	CSPSO	OLCS	CCS	ESCS
	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$
F_1	2.52E-28 \pm 2.31E-28	1.31E-26 \pm 1.05E-26	0.00E+00 \pm 0.00E+00	2.02E-30 \pm 1.01E-29
F_2	3.28E-11 \pm 9.03E-11	5.80E-02 \pm 5.16E-02	5.34E-09 \pm 7.74E-09	2.08E-09 \pm 4.88E-09
F_3	7.44E+05 \pm 6.43E+05	2.83E+06 \pm 7.55E+05	8.89E+04 \pm 5.33E+04	1.27E+05 \pm 1.06E+05
F_4	8.41E+01 \pm 1.05E+02	2.01E+03 \pm 8.15E+02	3.20E+01 \pm 9.74E+01	1.88E+01 \pm 1.91E+01
F_5	2.89E+03 \pm 7.82E+02	2.54E+03 \pm 5.86E+02	6.84E+02 \pm 6.65E+02	5.83E+02 \pm 4.32E+02
F_6	2.31E+00 \pm 4.79E+00	2.52E+01 \pm 2.04E+01	9.31E-01 \pm 2.58E+00	4.78E-01 \pm 1.32E+00
F_7	7.40E-03 \pm 2.20E-15	3.43E-04 \pm 4.42E-04	5.71E-03 \pm 6.58E-03	1.66E-02 \pm 1.32E-02
F_8	2.09E+01 \pm 5.49E-02	2.09E+01 \pm 5.75E-02	2.09E+01 \pm 7.18E-02	2.09E+01 \pm 1.04E-01
F_9	1.55E+02 \pm 2.47E+01	3.56E+01 \pm 6.56E+00	1.62E+01 \pm 5.19E+00	2.05E+01 \pm 5.97E+00
F_{10}	2.57E+02 \pm 6.81E+01	1.52E+02 \pm 3.48E+01	6.14E+01 \pm 1.48E+01	5.81E+01 \pm 1.22E+01

表3 不同反向学习模型的ESCS性能

Tab.3 Performance of ESCS by different learning models

函数	ESCS ₆	ESCS ₇	ESCS ₉	ESCS
	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$	$E_{avg} \pm E_{sd}$
F_1	0.00E+00 \pm 0.00E+00	0.00E+00 \pm 0.00E+00	2.02E-30 \pm 1.01E-29	2.02E-30 \pm 1.01E-29
F_2	4.28E-07 \pm 6.36E-07	2.94E-07 \pm 4.37E-07	3.81E-07 \pm 4.88E-09	2.08E-09 \pm 4.88E-09
F_3	1.12E+05 \pm 6.80E+04	1.55E+05 \pm 8.42E+04	1.43E+05 \pm 1.06E+05	1.27E+05 \pm 1.06E+05
F_4	8.66E+01 \pm 7.75E+01	8.31E+01 \pm 8.24E+01	1.29E+02 \pm 1.91E+01	1.88E+01 \pm 1.91E+01
F_5	1.03E+03 \pm 5.44E+02	9.91E+02 \pm 4.97E+02	8.74E+02 \pm 4.32E+02	5.83E+02 \pm 4.32E+02
F_6	3.51E+00 \pm 2.49E+00	4.48E+00 \pm 3.91E+00	4.89E+00 \pm 1.32E+00	4.78E-01 \pm 1.32E+00
F_7	1.38E-02 \pm 1.69E-02	1.31E-02 \pm 1.24E-02	1.18E-02 \pm 1.32E-02	1.66E-02 \pm 1.32E-02
F_8	2.09E+01 \pm 7.14E-02	2.09E+01 \pm 1.06E-01	2.09E+01 \pm 1.04E-01	2.09E+01 \pm 1.04E-01
F_9	2.31E+01 \pm 6.74E+00	2.39E+01 \pm 7.10E+00	2.13E+01 \pm 5.97E+00	2.05E+01 \pm 5.97E+00
F_{10}	6.24E+01 \pm 1.41E+01	6.78E+01 \pm 1.33E+01	6.13E+01 \pm 1.22E+01	5.81E+01 \pm 1.22E+01

解部分函数时获得较好性能.例如,ESCS₆算法在 F_1 和 F_3 上获得最佳性能;ESCS₇算法在 F_1 上获得最优值;ESCS₉算法求解 F_7 获得较好适应值;ESCS算法在较多的函数上获得较好解.同时,根据表4中各算法的平均排序可知,基于式(8)的ESCS算法具有较小的平均排序,这说明采用反向学习模型总体最优.

表4 4个不同反向学习模型的Friedman检验结果

Tab.4 Friedman results of the four algorithms by different learning models

函数	ESCS ₆	ESCS ₇	ESCS ₉	ESCS
排序值	2.75	2.75	2.85	1.65

4 结论

笔者提出一种基于精英反向学习和单纯形交叉以及发现概率自动控制相结合的布谷鸟搜索算法.仿真实验验证了提出的策略总体上能够改善算法的求精能力和收敛速度,从而也验证提出的策略是有效和可行的.下一步工作将提出的算法

应用于实际工程优化问题.

参考文献:

- [1] KARABODA D. An idea based on honey bee swarm for numerical optimization [R]. Kayseri, Turkey: Computer Engineering Department, Erciyes University, 2005.
- [2] YANG X. Nature-inspired metaheuristic algorithms [M]. Frome, UK: Luniver Press, 2008.
- [3] HE S, WU Q, SAUNDERS J. A novel group search optimizer inspired by animal behavioral ecology [C]// Proceedings of 2006 IEEE Congress on Evolutionary Computation. Vancouver, BC: IEEE, 2006: 1272-1278.
- [4] YANG X, DEB S. Cuckoo search via Lévy flight [C]// Proceedings of World Congress on Nature & Biologically Inspired Computing. Piscataway, NJ: IEEE, 2009: 210-214.
- [5] YANG X, DEB S. Engineering optimization by cuckoo search [J]. Int J Math Modelling Num Opt, 2010, 1(4): 330-343.
- [6] VALIAN E, MOHANNA S, TAVAKOLI S. Improved

- cuckoo search algorithm for global optimization [J]. Int J Communications and Information Technology, 2011, 1(1): 31–44.
- [7] WANG L, YIN Y, ZHONG Y. Cuckoo search with varied scaling factor [J]. Frontiers of computer science, 2015, 9(4): 623–635.
- [8] WANG L, ZHONG Y. Cuckoo search algorithm with chaotic maps [J]. Mathematical problems in engineering, 2015(1): 1–14.
- [9] 王李进, 尹义龙, 钟一文. 逐维改进的布谷鸟搜索算法 [J]. 软件学报, 2013, 24(11): 2687–2698.
- [10] 王李进, 钟一文, 尹义龙. 带外部存档的正交交叉布谷鸟搜索算法 [J]. 计算机研究与发展, 2015, 52(11): 2496–2507.
- [11] WANG F, HE X, LUO L, et al. Hybrid optimization algorithm of PSO and cuckoo search [C]//Proc of the 2nd Int Conf on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC). Piscataway, NJ: IEEE, 2011: 1172–1175.
- [12] LI X, WANG J, YIN M. Enhancing the performance of cuckoo search algorithm using orthogonal learning method [J]. Neural computing & applications, 2014, 24(6): 1233–1247.
- [13] WANG L, ZHONG Y. One position inheritance based cuckoo search algorithm [J]. Int J Computing Science and Mathematics, 2015, 6(6): 546–554.
- [14] 王晖. 基于一般反向学习的群体随机搜索算法框架 [J]. 南昌工程学院学报, 2012, 31(3): 1–6.
- [15] TIZHOOSH H. Opposition-based learning: a new scheme for machine intelligence [C]//Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce. Piscataway, NJ: IEEE, 2005, 1: 695–701.
- [16] TSUTSUI S, YAMAMURA M, HIGUCHI T. Multi-parent recombination with simplex crossover in real coded genetic algorithms [C]//Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation (GECCO'99). Orlando, Florida, USA: CECCO, 1999: 657–664.
- [17] NOMAN N, IBA H. Accelerating differential evolution using an adaptive local search [J]. IEEE transaction on evolutionary computation, 2008, 12(1): 107–125.
- [18] SUGANTHAN P, HANSEN N, LIANG J, et al. Problem definitions and evaluation criteria for the CEC2005 special session on real-parameter optimization [R]. Singapore, Singapore: Nanyang Technological University, 2005.

Elite Opposition-based Learning Based Simplex Crossover Cuckoo Search Algorithm

LIN Yaohua, WANG Lijin

(College of Computer And Information Science, Fujian Agriculture And Forestry University, Fuzhou 350002, China)

Abstract: Cuckoo search algorithm iteratively uses Lévy Flights random walk and Biased random walk to search for new individuals. In this paper, an enhanced cuckoo search was proposed, which employed elite opposition-based learning, simplex crossover and parameter control for the fraction probability. The elite opposition-based learning strategy was used to avoid the new individuals being homogeneous in the Lévy Flights random walk. The simplex crossover strategy was utilized to reduce the inefficiency of Biased random walk. The chaotic map was used to adaptively adjust the parameter pa to balance the exploration and the exploitation. The results of experiment showed the proposed strategies were overall effective, and make a great improvement on the performance of solution and convergence.

Key words: cuckoo search algorithm; simplex crossover; opposite learning; chaotic maps

考虑公差原则的三维公差分析方法研究

赵凤霞, 金少搏, 李纪峰

(郑州大学 机械工程学院, 河南 郑州 450001)

摘要: 针对目前公差分析中普遍存在未考虑公差原则影响的现象, 提出了一种考虑公差原则的三维公差建模与分析方法. 基于 SDT 理论和模态区间方法分别建立了要素采用独立原则、包容要求、最大实体要求、最小实体要求或可逆要求时的公差模型; 利用空间矢量表示三维尺寸链, 基于空间矢量环叠加原理推导出了三维公差分析时的封闭环尺寸及其变动计算模型; 以采用不同公差原则的工程实例进行了三维公差建模和公差分析, 给出了公差分析的步骤, 验证了本文所研究方法的有效性.

关键词: 公差原则; 公差分析; 公差建模; SDT 理论; 模态区间方法

中图分类号: TH124

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.007

0 引言

装配公差分析或称装配偏差分析, 即通过已知零部件的尺寸分布和公差, 考虑公差的累积和传播, 计算装配体的尺寸分布和装配公差的过程. 公差分析为解决零件尺寸公差和几何公差对装配性能的影响提供了一个有效的评价工具. 目前, 国内外学者在公差分析领域进行了诸多研究, 并取得了丰硕的成果. Chase 等^[1]和 Gao 等^[2]通过对隐式运动学方程进行线性化, 提出了一种基于矢量环模型的直接线性化公差分析方法, 该方法可以对二维和三维装配体进行分析. Davidson 等^[3]基于单纯形和面积坐标的概念提出了 T-Map 模型, 一个 T-Map 模型将特征公差带影射到一个多维参数空间区域, 该模型只能处理简单的公差链, 目前仍处于发展阶段. Desrochers 等^[4]将机器人运动学中的 Jacobian 模型和 SDT 模型相结合提出了 Jacobian-Torsor 模型. Li 等^[5]基于 SDT (small displacement torsor) 和误差传播理论进行了公差分析. 彭和平等^[6]基于矢量环装配模型进行了考虑几何公差的二维公差分析, 吴常林等^[7]和程彬彬等^[8]利用雅可比旋量理论建立了装配功能要求与零件公差之间的三维公差数学模型, 采用蒙特卡洛方法实现了装配体三维公差统计分析. 张为民等^[9]基于雅克比旋量法进行了实际工况的公差建模. 吴兆

强^[10]研究了小位移旋量的公差建模方法, 采用齐次坐标变换进行了三维公差分析. 但是, 目前的公差分析主要是研究尺寸公差分析, 或只对装配体中零件的尺寸公差和几何公差逐个进行零件装配分析, 而不考虑尺寸公差和几何公差之间相互关系及其对零件性能的影响. 实际上, 对同一个零件而言, 尺寸公差与几何公差是共存的, 二者之间存在一定的内在联系, 零件是否能装配, 并不单独取决于尺寸公差或者几何公差, 而是取决于尺寸公差和几何公差的综合作用. 因此, 在公差分析时必须考虑公差原则对装配性能的影响.

为了研究考虑公差原则的三维公差建模与分析方法, 笔者首先基于 SDT 理论和模态区间方法建立不同公差原则下的三维公差 SDT 模型, 然后基于空间矢量环叠加原理推导尺寸链的三维公差分析模型, 最后结合工程实例给出公差分析的步骤.

1 公差分析中的公差建模

公差建模是公差分析的基础. 笔者基于 SDT 理论和模态区间方法建立不同公差原则下的公差模型.

1.1 应用独立原则时的公差模型

应用独立原则时, 尺寸公差和几何公差是相互独立的, 二者相互无关. 尺寸公差只控制要素实际尺寸的变动量, 不控制几何误差; 而几何公差只

收稿日期: 2017-02-02; 修订日期: 2017-08-04

基金项目: 国家自然科学基金资助项目(51775515)

作者简介: 赵凤霞(1971—), 女, 河南扶沟人, 郑州大学教授, 博士, 主要从事精密测量技术方面的研究, E-mail: zfx-mail@163.com.

控制被测要素的几何误差,与实际尺寸无关.因此公差分析时应分别计入尺寸公差和几何公差,其中几何公差的尺寸为0.

尺寸公差和几何公差的特征由公差带来表征,实际要素可以在公差带内旋转或移动,这些旋转和移动可用SDT模型描述.设 ω 是旋转矢量; ε 是平动矢量; u, v, w 分别是沿 X, Y, Z 轴的平动量; $\Delta\alpha, \Delta\beta, \Delta\gamma$ 分别是沿 X, Y, Z 轴的旋转量,那么,公差带的SDT模型可定义为

$$\tau = \begin{bmatrix} \varepsilon \\ \omega \end{bmatrix} = \begin{bmatrix} u & v & w \\ \Delta\alpha & \Delta\beta & \Delta\gamma \end{bmatrix}. \quad (1)$$

几何公差带的形状主要有两平行平面、两平行直线、两同心圆、两同轴圆柱面、圆柱、圆、球、两平行曲线、两平行曲面等.如表1中的两平行平面公差带,平面长为 L_1 、宽为 L_2 ,两平行平面之间的距离为公差带大小 t .假设将坐标系 XYZ 置于两平行

平面中间.由新一代GPS的恒定度知^[11],平面有3个自由度,即绕 X 轴的转动 $\Delta\alpha$ 、绕 Y 轴的转动 $\Delta\beta$ 和沿 Z 轴的平动 w .因此,对两平行平面公差带建模只需确定这3个自由度的变动范围即可,由此可得其SDT模型及各矢量约束,各矢量约束采用模态区间法描述,见表1.同理,也可建立其他几何公差带的SDT模型,表1也给出了其他几种常用的几何公差带的SDT模型及其矢量约束.

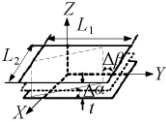
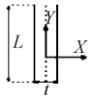
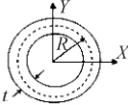
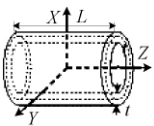
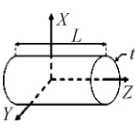
尺寸公差是允许尺寸的变动量,其公差带限制了实际偏差处于上极限偏差 T_U 和下极限偏差 T_L 区域内.根据尺寸公差的性质可知,尺寸公差带的SDT模型中只有平动矢量.假设平动发生在 X 方向上,则其SDT模型为

$$\tau = \begin{bmatrix} u & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (2)$$

其中, $u = [T_L, T_U]$.

表1 常用几何公差带的SDT模型及其约束

Tab.1 SDT model and its vector description of tolerance zones

公差带	公差带形状	SDT 模型	约束
两平行平面之间的区域		$\begin{bmatrix} 0 & 0 & w \\ \Delta\alpha & \Delta\beta & 0 \end{bmatrix}$	$w = \left[-\frac{t}{2}, \frac{t}{2}\right], \Delta\alpha = \left[-\frac{t}{L_1}, \frac{t}{L_1}\right], \Delta\beta = \left[-\frac{t}{L_2}, \frac{t}{L_2}\right]$ $ w + \left \frac{\Delta\alpha \cdot L_1}{2}\right + \left \frac{\Delta\beta \cdot L_2}{2}\right = \left[0, \frac{t}{2}\right]$
两平行直线之间的区域		$\begin{bmatrix} u & 0 & 0 \\ 0 & 0 & \Delta\gamma \end{bmatrix}$	$u = \left[-\frac{t}{2}, \frac{t}{2}\right], \Delta\gamma = \left[-\frac{t}{L}, \frac{t}{L}\right]$ $ u + \left \frac{\Delta\gamma \cdot L}{2}\right = \left[0, \frac{t}{2}\right]$
两同心圆之间的区域		$\begin{bmatrix} u & v & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$u = \left[-\frac{t}{2}, \frac{t}{2}\right], v = \left[-\frac{t}{2}, \frac{t}{2}\right], u^2 + v^2 = \left[0, \left(\frac{t}{2}\right)^2\right]$
两同轴圆柱面内的区域		$\begin{bmatrix} u & v & 0 \\ \Delta\alpha & \Delta\beta & 0 \end{bmatrix}$	$u = \left[-\frac{t}{2}, \frac{t}{2}\right], v = \left[-\frac{t}{2}, \frac{t}{2}\right], \Delta\alpha = \left[-\frac{t}{L}, \frac{t}{L}\right],$ $\Delta\beta = \left[-\frac{t}{L}, \frac{t}{L}\right]$ $\left(u + \frac{\Delta\alpha \cdot L}{2}\right)^2 + \left(v + \frac{\Delta\beta \cdot L}{2}\right)^2 = \left[0, \left(\frac{t}{2}\right)^2\right]$
圆柱面内的区域		$\begin{bmatrix} u & v & 0 \\ \Delta\alpha & \Delta\beta & 0 \end{bmatrix}$	$u = \left[-\frac{t}{2}, \frac{t}{2}\right], v = \left[-\frac{t}{2}, \frac{t}{2}\right], \Delta\alpha = \left[-\frac{t}{L}, \frac{t}{L}\right],$ $\Delta\beta = \left[-\frac{t}{L}, \frac{t}{L}\right]$ $\left(u + \frac{\Delta\alpha \cdot L}{2}\right)^2 + \left(v + \frac{\Delta\beta \cdot L}{2}\right)^2 = \left[0, \left(\frac{t}{2}\right)^2\right]$

1.2 应用包容要求时的公差模型

应用包容要求时,尺寸公差具有双重功能,既能控制实际尺寸,又能控制形状误差.尺寸要素对应的实际轮廓(实际尺寸和形状误差综合的结果)应遵守最大实体边界(MMB).因此,公差分析

时只需计入尺寸公差,而不考虑几何公差.尺寸公差的SDT模型同公式(2).

1.3 应用最大实体要求或最小实体要求时的公差模型

应用最大实体要求(MMR)的尺寸要素对应

的实际轮廓应遵守最大实体实效边界 (MMVB), 此时尺寸公差可以对几何公差进行补偿, 使几何公差带大小由 t 可增大到 $t+T$, T 为尺寸公差值. 当实际尺寸等于最小实体尺寸 (LMS) 时, 几何公差得到最大的补偿值, 几何公差带大小为 $t+T$. 因此, 应用 MMR 的尺寸要素可等效为尺寸等于 LMS 的理想形体 (公差值为 0) 与公差带大小等于 $t+T$ 的变动形体的叠加, 变动形体是尺寸为 0 的几何公差带, 其 SDT 模型与表 (1) 中对应相同形状模型相同, 但矢量约束中的 t 替换为 $t+T$. 为了计算方便, 应用 MMR 的尺寸要素进行公差分析时, 可将尺寸公差和几何公差合并为尺寸等于 LMS、具有 $t+T$ 公差带大小的综合公差.

如图 1 所示是圆柱轴应用 MMR 的示例, 公差分析时, 计入的尺寸等于 19.98 mm, 几何公差的公差带形状为圆柱, 由表 1 知, 其 SDT 模型为 $\begin{bmatrix} u & v & 0 \\ \Delta\alpha & \Delta\beta & 0 \end{bmatrix}$, 其中 $u = v = \left[-\frac{0.05}{2}, \frac{0.05}{2}\right]$, $\Delta\alpha = \Delta\beta = \left[-\frac{0.05}{100}, \frac{0.05}{100}\right]$.

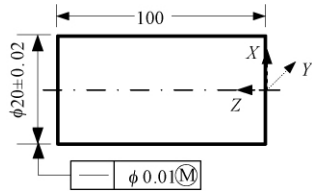


图 1 圆柱轴应用 MMR 示例

Fig. 1 Example of MMR for an cylindrical feature

应用最小实体要求 (LMR) 的尺寸要素对应的实际轮廓应遵守最小实体实效边界 (LMVB). 同理, 公差分析时, 将尺寸要素等效为尺寸等于 MMS 的理想形体 (公差为 0) 与公差带大小为 $t+T$ 的变动形体 (尺寸为 0) 的叠加.

1.4 应用可逆要求时的公差模型

可逆要求 (RPR) 是最大实体要求 (MMR) 或最小实体要求 (LMR) 的附加要求. MMR 或 LMR 附加了 RPR 后, 允许尺寸和几何公差之间相互补偿; 当尺寸要素具有理想形状或方向或位置时, 尺寸得到最大的补偿值. 因此, 与 1.3 节同理, 把尺寸要素等效为尺寸一定的理想形体与公差带大小为 $t+T$ 的变动形体的叠加, 变动形体的 SDT 模型与应用 MMR 和 LMR 的相同. 理想形体的尺寸与尺寸要素的形状和采用的相关要求有关, 如表 2 所示是笔者推导出的应用可逆要求时理想形体尺寸的计算公式.

表 2 应用可逆要求时理想形体尺寸的计算公式

Tab. 2 The formula of basic size of nominal shape when using reciprocity requirement.

相关要求	尺寸要素的形状	
	外尺寸 (轴)	内尺寸 (孔)
RPR 用于 MMR	MMS + t	MMS - t
RPR 用于 LMR	LMS - t	LMS + t

2 公差分析中的封闭环模型建立

三维公差分析即确定三维尺寸链中各个组成环矢量对封闭环的影响, 从而确定封闭环的公差.

在三维空间中, 每一个矢量的空间位置是由矢量长度 a_i 和其与三个坐标轴的夹角 $\alpha_i, \beta_i, \gamma_i$ 确定, 如图 2 所示, 即各矢量的三维坐标为

$$A_i = [a_i \cos \alpha_i \quad a_i \cos \beta_i \quad a_i \cos \gamma_i]. \quad (3)$$

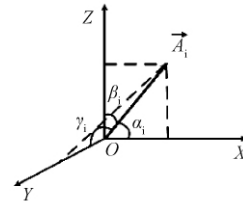


图 2 矢量的空间表示

Fig. 2 the space representation of vector

假设三维尺寸链由 4 个空间组成环矢量 (A_1, A_2, A_3, A_4) 组成, A_0 为封闭环. 由尺寸链的特点, 对各矢量环进行叠加可得封闭环尺寸的三维模型为

$$A_0 = A_1 + A_2 + A_3 + A_4 =$$

$$\begin{bmatrix} \cos \alpha_1 & \cos \alpha_2 & \cos \alpha_3 & \cos \alpha_4 \\ \cos \beta_1 & \cos \beta_2 & \cos \beta_3 & \cos \beta_4 \\ \cos \gamma_1 & \cos \gamma_2 & \cos \gamma_3 & \cos \gamma_4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}. \quad (4)$$

假设由于尺寸公差和/或几何公差的影响, 使得各矢量的尺寸变动为 Δa_i , 与各坐标轴的夹角变动为 $\Delta \alpha_i, \Delta \beta_i, \Delta \gamma_i$, 则各组成环矢量的三维模型变为

$$A_{iw} = \begin{bmatrix} (a_i + \Delta a_i) \cos(\alpha_i + \Delta \alpha_i) \\ (a_i + \Delta a_i) \cos(\beta_i + \Delta \beta_i) \\ (a_i + \Delta a_i) \cos(\gamma_i + \Delta \gamma_i) \end{bmatrix}^T. \quad (5)$$

由于 $\Delta \alpha_i \rightarrow 0, \Delta \beta_i \rightarrow 0, \Delta \gamma_i \rightarrow 0, \Delta a_i \rightarrow 0$, 可以认为 $\cos \Delta \alpha_i \approx 1, \cos \Delta \beta_i \approx 1, \cos \Delta \gamma_i \approx 1$, $\sin \Delta \alpha_i \approx \Delta \alpha_i, \sin \Delta \beta_i \approx \Delta \beta_i, \sin \Delta \gamma_i \approx \Delta \gamma_i$; 且 $\Delta a_i \Delta \alpha_i \approx 0, \Delta a_i \Delta \beta_i \approx 0, \Delta a_i \Delta \gamma_i \approx 0$.

则 $\cos(\alpha_i + \Delta \alpha_i) \approx \cos \alpha_i - \Delta a_i \sin \alpha_i$;

$$\cos(\beta_i + \Delta\beta_i) \approx \cos \beta_i - \Delta\beta_i \sin \beta_i;$$

$$\cos(\gamma_i + \Delta\gamma_i) \approx \cos \gamma_i - \Delta\gamma_i \sin \gamma_i.$$

由此可得封闭环的变动模型为

$$\Delta = A_{0b} - A_0 =$$

$$\begin{bmatrix} \cos \alpha_1 & \cos \beta_1 & \cos \gamma_1 \\ \cos \alpha_2 & \cos \beta_2 & \cos \gamma_2 \\ \cos \alpha_3 & \cos \beta_3 & \cos \gamma_3 \\ \cos \alpha_4 & \cos \beta_4 & \cos \gamma_4 \\ -a_1 \sin \alpha_1 & 0 & 0 \\ -a_2 \sin \alpha_2 & 0 & 0 \\ -a_3 \sin \alpha_3 & 0 & 0 \\ -a_4 \sin \alpha_4 & 0 & 0 \\ 0 & -a_1 \sin \beta_1 & 0 \\ 0 & -a_2 \sin \beta_2 & 0 \\ 0 & -a_3 \sin \beta_3 & 0 \\ 0 & -a_4 \sin \beta_4 & 0 \\ 0 & 0 & -a_1 \sin \gamma_1 \\ 0 & 0 & -a_2 \sin \gamma_2 \\ 0 & 0 & -a_3 \sin \gamma_3 \\ 0 & 0 & -a_4 \sin \gamma_4 \end{bmatrix} \cdot \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \Delta a_3 \\ \Delta a_4 \\ \Delta \alpha_1 \\ \Delta \alpha_2 \\ \Delta \alpha_3 \\ \Delta \alpha_4 \\ \Delta \beta_1 \\ \Delta \beta_2 \\ \Delta \beta_3 \\ \Delta \beta_4 \\ \Delta \gamma_1 \\ \Delta \gamma_2 \\ \Delta \gamma_3 \\ \Delta \gamma_4 \end{bmatrix}.$$

(6)

式(6)可扩展到 n 个组成环矢量情况.

3 公差分析步骤及工程实例

公差分析的步骤:

(1) 首先对装配体上的所有零件进行分析,

识别对保证功能有影响的所有要素.

图3所示是一个机床尾座的装配简图及组成零件图. 该装配体由顶尖(图3中的零件1)、立柱(图3中的零件2)、底座(图3中的零件3)组成, 其装配功能要求是保证顶尖端点在垂直方向上的尺寸 FR . 顶尖与立柱的孔轴配合为 $\phi 20 \frac{H6}{n5}$, 底座与立柱的配合为平面配合, 其余尺寸如图示. 根据工作条件可知, 顶尖上的功能要素有: 圆锥轴线 F_1 、 $\phi 20$ 圆柱的轴线 F_2 、 $\phi 20$ 圆柱面 F_3 . 底座和立柱的功能要素见图3.

(2) 对各功能要素建立坐标系.

(3) 根据零件间的装配关系基于图论建立零件功能要素间的装配关系图.

图4所示是图3示例中各零件功能要素间的装配关系图. 图中, 每个零件内部要素之间的尺寸和方位关系在内圆中示出, 零件之间的要素关系在外圆外分别实线或虚线连接, 实线表示两零件要素之间的装配关系, 虚线表示设计要求. 各链环的矢量长度分别用 $a_i (i = 1, 2, 3, \dots)$ 表示.

(4) 根据各功能要素所采用的公差原则按本文第1章所述内容进行公差建模.

如顶尖零件中, 要素 F_1 对要素 F_2 有同轴度要求, 且采用了独立原则, 由1.1节知, 公差分析时计入尺寸公差和几何公差两部分. 其中, 尺寸及公差为0; 几何公差的尺寸为0, 用记号 a_1 表示, 同轴度的公差带形状为圆柱, 由表1可知其SDT

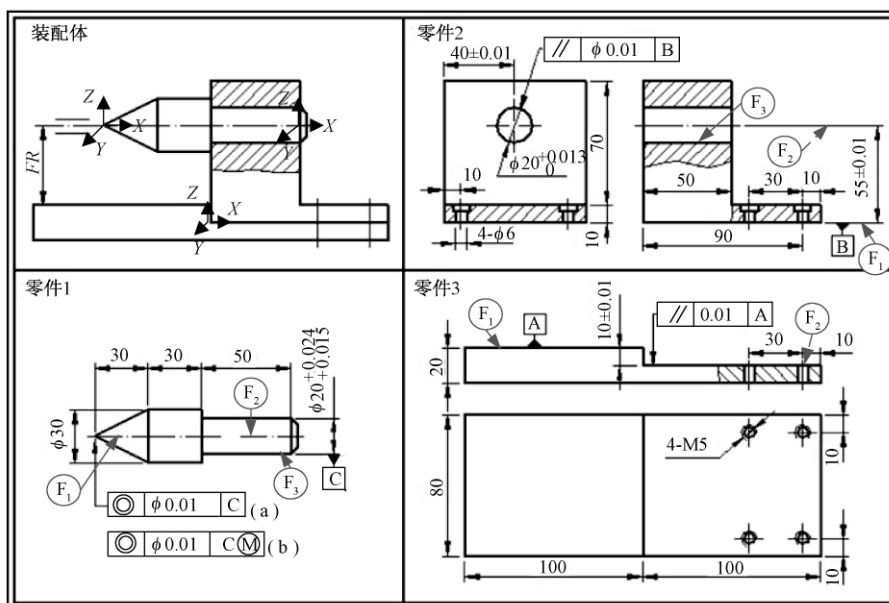


图3 机床尾座的装配简图及组成零件图

Fig. 3 Assembly drawing of machine tailstock and its parts drawing

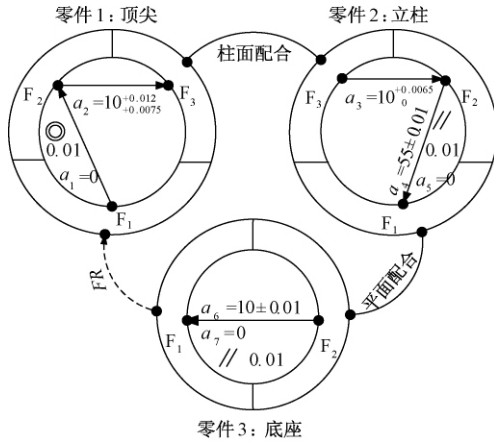


图4 功能要素间的装配关系图

Fig. 4 The assembly graph of effective features

模型为 $\begin{bmatrix} 0 & v & w \\ 0 & \Delta\beta & \Delta\gamma \end{bmatrix}$, $v = w = \begin{bmatrix} -\frac{0.01}{2}, \frac{0.01}{2} \end{bmatrix}$,

$$\Delta\beta = \Delta\gamma = \begin{bmatrix} -\frac{0.01}{30}, \frac{0.01}{30} \end{bmatrix}.$$

要素 F_2 为基准 C, 当同轴度采用图 3 中 (a) 方式标注时, 即 C 后面无 (M) , 根据基准的建立准则可知^[12], 此时基准 C 相当于采用了包容要求. 由 1.2 节知, 公差分析时, 只计入 F_2 到 F_3 的尺寸及公差 $a_2 = \phi 10^{+0.012}_{-0.0075}$, 其 SDT 模型及其约束见表 3. 当同轴度采用 (b) 方式标注时, C 后面有 (M) , 即基准 C 采用了最大实体要求. 由 1.3 节知, 公差分析时, F_2 到 F_3 的尺寸及公差为 $a'_2 = \phi 10.0075 \pm 0.00225$. 同理, 也可得立柱和底座各功能要素的公差模型, 结果见表 3.

(5) 用公式 (4) 和公式 (6) 计算得到封闭环的尺寸及其变动.

表3 各矢量环的 SDT 模型及其约束

Tab. 3 The SDT modals and its vector description of the vectors

各环尺寸及其 $\alpha_i, \beta_i, \gamma_i$	SDT 模型	约束
$a_1 = 0$ $\alpha_1 = 90^\circ, \beta_1 = 90^\circ, \gamma_1 = 0^\circ$	$\begin{bmatrix} 0 & v & w \\ 0 & \Delta\beta & \Delta\gamma \end{bmatrix}$	$v = w = \begin{bmatrix} -\frac{0.01}{2}, \frac{0.01}{2} \end{bmatrix}$ $\Delta\beta = \Delta\gamma = \begin{bmatrix} -\frac{0.01}{30}, \frac{0.01}{30} \end{bmatrix}$
$a_2 = 10^{+0.012}_{-0.0075}$ $\alpha_2 = 90^\circ, \beta_2 = 90^\circ, \gamma_2 = 180^\circ$	$\begin{bmatrix} 0 & 0 & w \\ 0 & 0 & 0 \end{bmatrix}$	$w = [0.0075, 0.012]$
$a'_2 = 10.0075 \pm 0.00225$ $\alpha_2 = 90^\circ, \beta_2 = 90^\circ, \gamma_2 = 180^\circ$	$\begin{bmatrix} 0 & 0 & w \\ 0 & 0 & 0 \end{bmatrix}$	$w = [-0.00225, 0.00225]$
$a_3 = 10^{+0.0065}_{-0.001}$ $\alpha_3 = 90^\circ, \beta_3 = 90^\circ, \gamma_3 = 0^\circ$	$\begin{bmatrix} 0 & 0 & w \\ 0 & 0 & 0 \end{bmatrix}$	$w = [0, 0.0065]$
$a_4 = 55 \pm 0.01$ $\alpha_4 = 90^\circ, \beta_4 = 90^\circ, \gamma_4 = 180^\circ$	$\begin{bmatrix} 0 & 0 & w \\ 0 & 0 & 0 \end{bmatrix}$	$w = [-0.01, 0.01]$
$a_5 = 0$ $\alpha_5 = 90^\circ, \beta_5 = 90^\circ, \gamma_5 = 180^\circ$	$\begin{bmatrix} 0 & v & w \\ 0 & \Delta\beta & \Delta\gamma \end{bmatrix}$	$v = w = \begin{bmatrix} -\frac{0.01}{2}, \frac{0.01}{2} \end{bmatrix}$ $\Delta\beta = \Delta\gamma = \begin{bmatrix} -\frac{0.01}{50}, \frac{0.01}{50} \end{bmatrix}$
$a_6 = 10 \pm 0.01$ $\alpha_6 = 90^\circ, \beta_6 = 90^\circ, \gamma_6 = 0^\circ$	$\begin{bmatrix} 0 & 0 & w \\ 0 & 0 & 0 \end{bmatrix}$	$w = [-0.01, 0.01]$
$a_7 = 0$ $\alpha_7 = 90^\circ, \beta_7 = 90^\circ, \gamma_7 = 0^\circ$	$\begin{bmatrix} 0 & 0 & w \\ \Delta\alpha & \Delta\beta & 0 \end{bmatrix}$	$w = \begin{bmatrix} -\frac{0.01}{2}, \frac{0.01}{2} \end{bmatrix}$ $\Delta\alpha = \Delta\beta = \begin{bmatrix} -\frac{0.01}{100}, \frac{0.01}{100} \end{bmatrix}$

将表 3 中的内容分别代入公式 (4) 和公式 (6), 用模态区间运算法则进行计算^[13], 得到 FR 在 Z 方向的尺寸为 45 mm. 当同轴度采用方式 (a) 标注时, 变动为 $[-0.047, 0.034]$; 当同轴度采用方式 (b) 标注时, 变动为 $[-0.0475, 0.036]$. 对比计算结果可以看出, 采用的公差原

则不同, 公差分析结果亦不相同, 因此在实际工作中不能忽视公差原则的影响.

根据表 3 中 SDT 模型各矢量变动的特点, 也可假设各环的分布状态 (如假设 A_2, A_3 中的矢量呈均匀分布, 其他各环中的矢量呈正态分布) 采用蒙特卡洛方法或其他方法进行分析, 笔者不再

赘述.

4 结论

笔者研究了考虑公差原则的三维公差分析问题,得到了在公差分析时应用不同公差原则的尺寸公差和几何公差的计入方法,并基于 SDT 理论和模态区间方法分别建立了要素应用不同公差原则时的公差模型;基于空间矢量环叠加原理推导出了三维公差分析时的封闭环尺寸及其变动计算模型;结合工程实例给出了公差分析的具体步骤,并进行了公差分析的实例应用.研究表明,公差原则对装配精度的影响较大,在实际工作中是不能忽视的,考虑公差原则的三维公差分析结果更为精确.笔者的研究为准确且符合工程实际进行三维公差分析提供了技术支撑.

参考文献:

- [1] CHASE K W, GAO J, MAGLEBY S P. General 2D tolerance analysis of mechanical assemblies with small kinematic adjustments [J]. J Des Manuf, 1995, 5 (4): 263 - 274.
- [2] GAO J, CHASE K W, MAGLEBY S P. Generalized 3D tolerance analysis of mechanical assemblies with small kinematic adjustments [J]. IIE transactions, 1998, 30(4): 367 - 377.
- [3] DAVIDSON J K, MUJEZINOVIC A, SHAH J J. A new mathematical model for geometric tolerances as applied to round faces [J]. J Mech Des, 2002, 124: 609 - 621.
- [4] DESROCHERS A, GHIE W, LAPERRIERE L. Application of a unified Jacobian-torsor model for tolerance analysis [J]. J Comput Inf Sci Eng, 2003, 3 (1): 2 - 14.
- [5] LI H, ZHU H P, LI P G, et al. Tolerance analysis of mechanical assemblies based on small displacement torsor and deviation propagation theories [J]. Int J Adv Manuf Technol, 2014, 72: 89 - 99.
- [6] 彭和平, 刘晓军. 考虑形位公差的二维装配公差分析 [J]. 机械传动, 2008, 32(3): 75 - 82.
- [7] 吴常林, 黄美发, 程彬彬, 等. 基于雅可比旋量模型的装配体公差统计分析 [J]. 机械设计与研究, 2015, 31(6): 97 - 100.
- [8] 程彬彬, 王志越, 黄美发, 等. 装配体三维公差建模及分析研究 [J]. 机械设计与制造, 2016(4): 76 - 79.
- [9] 张为民, 陈灿, 李鹏忠, 等. 基于雅可比旋量法的实际工况公差建模 [J]. 计算机集成制造系统, 2011, 17(1): 77 - 83.
- [10] 吴兆强. 基于小位移旋量的公差建模及公差分析 [J]. 机械设计与制造, 2010(1): 205 - 207.
- [11] 方东阳, 李秀明, 张琳娜, 等. 基于 GPS 的形位公差项目分析和设计 [J]. 郑州大学学报(工学版), 2005, 26(3): 73 - 75.
- [12] Geometrical product specifications (GPS) -geometrical tolerancing-datums and datum systems: ISO 5459: 2011 (E) [S].
- [13] SAINZ M A, ARMENGO J, CALM R, et al. Modal interval analysis: new tools for numerical information [M]. Berlin: Springer, 2014.

Research on the Method for 3D Tolerance Analysis Considering Tolerance Principle

ZHAO Fengxia, JIN Shaobo, LI Jifeng

(School of Mechanical Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: A method of considering tolerance principle for three dimensional tolerance analysis was put forward. Based on small displacement torsor (SDT) theory and modal interval arithmetic, the tolerance models of size tolerance and geometrical tolerance of the feature of size apply independent principle, envelope requirement, maximum material requirement, least material requirement or reciprocity requirement, were established respectively. By using the space vector to represent 3D dimension chain, a mathematical model was built to calculate the closed loop tolerance based on space vector loop stack principle. The application of the proposed method was illustrated through presenting an example, the tolerance analysis steps were given, and the availability of the proposed method was proved successfully.

Key words: tolerance principle; tolerance analysis; tolerance modeling; small displacement torsor theory; modal interval arithmetic

计及线路损耗的峰谷分时电价优化模型

赵国生, 詹天乐, 李 博

(郑州大学 电气工程学院, 河南 郑州 450001)

摘 要: 针对现有的峰谷分时电价优化模型建立过程中没有考虑线路损耗的问题, 基于线路损耗和负荷波动之间的关系, 利用电量电价弹性矩阵, 以峰谷差和线路损耗的最小化为目标函数建立峰谷分时电价优化模型, 并通过理想点法将其转化为单目标模型后求解。仿真证明, 计及线路损耗的峰谷分时电价优化模型很好地实现了削峰填谷、减少线路损耗的目的。

关键词: 峰谷分时电价; 线路损耗; 用户电量电价弹性矩阵; 多目标优化模型; 理想点法

中图分类号: TM73

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2016.06.022

0 引言

20世纪80年代以来, 峰谷分时电价在我国逐步开始实行。作为电力需求侧管理^[1]的所有经济手段实施范围最广的一种, 关于其制定策略的研究一直是电力需求侧管理研究的热点。目前国内关于峰谷分时电价制定策略的研究多数集中在通过用户对电价的需求响应矩阵或电价反应曲线, 从减少电网负荷波动、提高电力系统经济效益的角度出发, 建立优化模型方面。文献[2]基于用户对电价的响应矩阵, 以峰谷差最小为目标函数建立峰谷分时电价优化模型并利用遗传算法进行求解; 文献[3]分析了用户用电响应特点, 基于多智能体技术建立峰谷分时电价模型; 文献[4]基于用户电价反应曲线, 以减少电网峰负荷和峰谷差为目标函数建立峰谷分时电价优化模型并进行模糊求解; 文献[5]从供电公司的利益出发, 以供电公司收益最大化为目标函数建立峰谷分时电价模型。目前我国全社会用电量已超过5万亿千瓦时, 据统计有超过6%^[6]的电能损耗在输电线路路上, 因此减少线路损耗对电网有重要的经济意义, 但是以上文献在建立峰谷分时电价优化模型的过程中均未考虑线路损耗的影响。

文献[7]分析了负荷曲线和线路损耗的关系, 通过仿真说明了峰谷分时电价的实施能有效地减少线路损耗, 但是并没有给出峰谷分时电价

优化模型建立的方法。笔者在此基础上以更简洁的方式证明了线路损耗和线路负荷曲线的关系, 从减少线路损耗的角度出发, 建立峰谷分时电价优化模型, 具有一定的实际意义。

1 用户电量电价弹性矩阵

用户对电力需求的价格弹性指的是由于电价的变动, 用户的用电量相应变化的程度。峰谷分时电价将一天分成峰、平、谷三个时段, 不同时段用电价格不同。用户会将电价较高时段的负荷转移到其他时段, 因此用户在某个时段的用电量不仅与本时段的电价有关, 还会受到其他时段电价的影响。为表示用户各时段用电量与各时段电价之间的关系, 一般利用用户电量电价弹性矩阵。

电量电价弹性矩阵由自弹性系数和交叉弹性系数组成, 分别表示本时段电价和其他时段电价对该时段负荷的影响, 于是各个时段负荷和电价之间的关系可以用下式表示:

$$\begin{bmatrix} \frac{Q'_1 - Q_1}{Q_1} \\ \frac{Q'_2 - Q_2}{Q_2} \\ \frac{Q'_3 - Q_3}{Q_3} \end{bmatrix} = \mathbf{E} \begin{bmatrix} \frac{p_1 - p_0}{p_0} \\ \frac{p_2 - p_0}{p_0} \\ \frac{p_3 - p_0}{p_0} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} \frac{p_1 - p_0}{p_0} \\ \frac{p_2 - p_0}{p_0} \\ \frac{p_3 - p_0}{p_0} \end{bmatrix}, \quad (1)$$

式中: \mathbf{E} 为电量电价弹性矩阵; e_{ii} 、 e_{ij} 分别表示自

收稿日期: 2016-06-30; 修订日期: 2016-09-19

基金项目: 国家自然科学基金资助项目(51307152)

作者简介: 赵国生(1965—), 男, 河南南阳人, 郑州大学副教授, 博士, 主要从事电能质量方面的研究, E-mail: zgs410@zzu.edu.cn.

弹性系数和交叉弹性系数; p_0 为实行峰谷电价前的统一电价; p_1 、 p_2 、 p_3 分别表示实行峰谷分时电价后峰、平、谷时段的电价; $Q_1(\theta_i)$ 、 $Q_2(\theta_i)$ 、 $Q_3(\theta_i)$ 分别表示实行峰谷分时电价前(后)处于实施峰谷电价后对应峰平谷时段的总负荷; 于是实行峰谷分时电价后峰平谷时段的总负荷为

$$Q'_i = Q_i \left(1 + \sum_{j=1}^3 e_{ij} \frac{p_j - p_0}{p_0} \right), i = 1, 2, 3. \quad (2)$$

按日负荷曲线取 24 个时刻的负荷考虑, 设 T_1 表示峰时段, T_2 表示平时段, T_3 表示谷时段, 执行峰谷分时电价后在峰、平、谷时段内时刻 t 负荷为

$$q'_i = q_i \left(1 + \sum_{j=1}^3 e_{ij} \frac{p_j - p_0}{p_0} \right), i = 1, 2, 3; t \in T_i, \quad (3)$$

式中: q_i 表示峰谷分时电价执行前时刻 t 的负荷大小; q'_i 表示峰谷分时电价执行后时刻 t 的负荷, $t \in T_i$; $i = 1, 2, 3$.

2 线路负荷波动对线路损耗的影响

电力系统中输电线路损耗包括可变损耗和不变损耗. 根据输电线路等值模型可知, 其中不变损耗与线路负荷没有关系, 而输电线路中电能损耗主要为可变损耗 $\psi(t)$.

$$\psi(t) = \int_0^T 3I^2(t) R \times 10^{-3} dt = \int_0^T \left(\frac{p(t)}{u \cos \theta} \right)^2 R \times 10^{-3} dt, \quad (4)$$

式中: $I(t)$ 表示输电线路相电流; T 表示用户负荷的运行周期(日、月、季、年等); U 表示用户负荷节点的线电压, kV; $\cos \theta$ 表示用户负荷的功率因数; $p(t)$ 表示用户负荷的有功功率, kW; R 表示输电线路等效电阻, Ω .

正常运行情况下, 系统各节点电压基本保持不变. 为了研究的方便, 可以假设系统各节点电压为常数, 同时认为峰谷分时电价只改变用户用电负荷时间上的分配, 不改变负荷的种类和性质, 即可以假设负荷的功率因数也是一定的. 由于实施峰谷分时电价以后, 用户总负荷一般变化很小, 本文证明在总负荷一定的情况下, 负荷波动越小线路损耗越小, 负荷为恒定值时线路损耗最小. 设 $\int_0^T p(t) dt = P$, 于是下式成立:

$$T \int_0^T p^2(t) dt - P^2 = T \int_0^T \left(p(t) - \frac{P}{T} \right)^2 dt. \quad (5)$$

根据概率论可知, 对于连续型随机变量 x , 若其概率密度函数为 $f(x)$, $y = y(x)$ 为 x 的函数, 则 y 的方差 σ^2 反应其在期望 $E(y(x))$ 附近波动的

程度, 也即曲线波动的程度:

$$\sigma^2 = \int_{-\infty}^{+\infty} (y(x) - E(y(x)))^2 f(x) dx. \quad (6)$$

若将时间 t 看成连续型随机变量, 则其概率密度函数在一个周期 T 内为 $1/T$. 负荷 $p(t)$ 为 t 的函数, 根据上式有 $p(t)$ 的方差 σ^2 为

$$\sigma^2 = \frac{1}{T} \int_0^T \left(p(t) - \frac{\int_0^T p(t) dt}{T} \right)^2 dt. \quad (7)$$

当负荷为恒定值 P/T 时, 线路损耗为

$$\psi_0(t) = \int_0^T \left(\frac{p(t)}{u \cos \theta} \right)^2 R \times 10^{-3} dt = \frac{R \times 10^{-3}}{(u \cos \theta)^2} \int_0^T \left(\frac{P}{T} \right)^2 dt. \quad (8)$$

对于一般的负荷曲线, 线路损耗为

$$\psi(t) = \int_0^T \left(\frac{p(t)}{u \cos \theta} \right)^2 R \times 10^{-3} dt = \frac{R \times 10^{-3}}{(u \cos \theta)^2} \int_0^T (p(t))^2 dt. \quad (9)$$

于是当且仅当 $p(t)$ 为恒定值 P/T 时, 二者比值为

$$\frac{\psi(t)}{\psi_0(t)} = \frac{T \int_0^T p^2(t) dt}{P^2} = 1 + \frac{T^2}{P^2} \sigma^2 \geq 1. \quad (10)$$

由此可见, 在用户总负荷一定的情况下, 当负荷为恒定值时线路损耗最小. 负荷波动越小, σ^2 就越小, 线路损耗就越小.

假设 $p_0(t)$ 、 $p_1(t)$ 分别是实施峰谷分时电价前和实施后相同运行周期 T 内对应的线路负荷曲线. 峰谷分时电价能够削峰填谷, 减少线路负荷波动, 因此能够减少线路损耗. 为了更好地衡量峰谷分时电价减少线路损耗的效果, 定义线路损耗比 ω 为峰谷分时电价实施后和实施前线路损耗比:

$$\omega = \frac{\psi_1(t)}{\psi_0(t)} = \frac{\int_0^T p_1^2(t) dt}{\int_0^T p_0^2(t) dt}. \quad (11)$$

如果以负荷曲线上一天 24 个时刻的负荷来代表线路负荷, 将上式分子和分母积分离散化表示后, ω 的表达式为

$$\omega = \frac{\sum_{t=1}^{24} p_1^2(t)}{\sum_{t=1}^{24} p_0^2(t)}. \quad (12)$$

3 峰谷分时电价优化模型及其求解

3.1 优化模型目标函数

实行峰谷分时电价的主要追求目标是为了削峰填谷减少负荷波动, 同时应该考虑提高电网经

济效益,减少线路上电能损耗.笔者以一天24个时刻的负荷代表负荷曲线,则目标函数可以如下表示为.

(1) 用户负荷峰谷差 ν 最小:

$$\min \nu = \min(\max_{t \in T_i} (q_t) - \min_{t \in T_j} (q_t)); i, j \in 1, 2, 3 \quad (13)$$

(2) 线路损耗比最小:

$$\min \omega = \min \sum_{i=1}^3 \sum_{t \in T_i} (q_t)^2 / \sum_{t=1}^{24} (q_t)^2. \quad (14)$$

3.2 优化模型约束条件

(1) 用户利益约束. 峰谷分时电价实行以后,为了得到用户的支持,用户需要在经济上得到实惠,因此用户平均电价应该减少:

$$p'_0 = \frac{p_1 Q'_1 + p_2 Q'_2 + p_3 Q'_3}{Q'_1 + Q'_2 + Q'_3} < p_0, \quad (15)$$

式中, p'_0 代表实施峰谷电价以后用户平均电价.

(2) 供电公司利益约束. 作为峰谷分时电价的实施方,供电公司实行峰谷分时电价后的利润也要得到保证. 实施峰谷分时电价以后,节省了投资,减少了供电公司运行成本,因此供电公司的总售电收入可以略有减少:

$$\begin{cases} P_0 = (p_0 - p_m)(Q_1 + Q_2 + Q_3); \\ P'_0 = p_1 Q'_1 + p_2 Q'_2 + p_3 Q'_3 - p_m(Q'_1 + Q'_2 + Q'_3); \\ (1 - M)P_0 \leq P'_0, \end{cases} \quad (16)$$

式中, P_0 和 P'_0 分别代表实施峰谷分时电价前、后供电公司一天的收入; p_m 表示当地平均上网电价; M 表示供电公司售电收入允许减少的最大比例.

(3) 用户总负荷约束. 实行峰谷分时电价以后,为了用户自身生产和生活的需要,用户的总负荷应保持基本不变,则

$$1 - \delta \leq \frac{Q'_1 + Q'_2 + Q'_3}{Q_1 + Q_2 + Q_3} \leq 1 + \delta, \quad (17)$$

式中, δ 表示实施峰谷分时电价以后允许用户总负荷变动的最大比例.

(4) 峰平谷电价约束. 峰平谷电价太接近起不到削峰填谷的效果,反之则会严重影响用户的生活和生产习惯,同时也会损害供电公司的利益,不利于峰谷分时电价的实施. 根据目前各省执行峰谷分时电价的实际执行情况设定以下约束:

$$\begin{cases} 1.4 \leq p_1/p_0 \leq 1.8; \\ 0.9 \leq p_2/p_0 \leq 1.1; \\ 0.5 \leq p_3/p_0 \leq 0.7. \end{cases} \quad (18)$$

于是,优化模型便可用式(13)、(14)表示的目标函数和式(15)~(18)表示的目标函数构成.

3.3 优化模型求解

峰谷分时电价优化模型是个非线性的多目标优化模型,一般没有绝对最优解,通常的做法是将其转化为单目标问题后求解. 在多目标优化模型的常见解法中,考虑到理想点法^[8]几何意义明显,并且能够避免取权重的主观性,笔者将其作为本文模型的解法. 理想点法比较经典的做法是将目标函数组成目标点,各个目标函数看成目标点的一个坐标,然后分别求出单个目标的最优值组成最优点,最后利用目标点和最优点之间的欧式距离的大小来衡量解的最优化程度.

具体求解过程如下: 单个目标的最优值可以利用 matlab 中的 fmincon 函数解决. 首先分别求出单个目标的最大值和最小值 ν_{\max} 、 ν_{\min} 、 ω_{\max} 、 ω_{\min} , 由于两个目标函数数值差别较大,先对它们做归一化处理:

$$\bar{\nu} = \frac{\nu - \nu_{\min}}{\nu_{\max} - \nu_{\min}}. \quad (19)$$

$$\bar{\omega} = \frac{\omega - \omega_{\min}}{\omega_{\max} - \omega_{\min}}. \quad (20)$$

于是 $\bar{\nu}$ 、 $\bar{\omega}$ 均属于 $[0, 1]$. $\bar{\nu}$ 、 $\bar{\omega}$ 能反映目标函数的最优化程度,越接近 0 代表最优化程度越高,因此目标点可以取 $(\bar{\nu}, \bar{\omega})$. $\bar{\nu}$ 、 $\bar{\omega}$ 最理想值为 0, 因此最优点取 $(0, 0)$. 根据理想点法,最终的目标函数 f 可以取目标点和最优点的欧氏距离,如下式所示:

$$\min f = \min \sqrt{\bar{\nu}^2 + \bar{\omega}^2}. \quad (21)$$

于是,最终的单目标优化模型可以由式(21)表示的目标函数和式(15)~(18)表示的约束条件构成.

4 算例分析

以文献[9]中安徽 10 kV 大工业的负荷数据为例,如表 1 所示.

表 1 典型日负荷数据
Tab. 1 typical daily load data

时刻	负荷/ MW	时刻	负荷/ MW	时刻	负荷/ MW
1:00	261	9:00	386	17:00	338
2:00	243	10:00	402	18:00	356
3:00	224	11:00	337	19:00	393
4:00	202	12:00	358	20:00	409
5:00	229	13:00	340	21:00	447
6:00	247	14:00	332	22:00	371
7:00	273	15:00	357	23:00	330
8:00	315	16:00	379	24:00	304

实行峰谷分时电价前的安徽 10 kV 大工业统一电价 p_0 为 0.602 0 元/(kW·h), 平均上网电价 p_m 为 0.4 元/(kW·h). 实行峰谷分时电价后, 根据该大工业负荷的特点将峰、平、谷时段划分: 高峰时段为 9:00 ~ 11:00、16:00 ~ 17:00、18:00 ~ 23:00; 低谷时段为 24:00 ~ 9:00; 其余时段为平时段. 根据文献[10]设供电公司收入减少最大允许值 $M=2\%$. 为了保证峰谷分时电价实施以后用户总负荷保持基本不变, 设用户总负荷允许最大变动 $\delta=2\%$. 电量电价弹性矩阵与用户类型、当地经济情况等因素都有关系. 笔者由于数据不足无法得到当地准确的电量电价弹性矩阵, 文献[11]比较科学地分析了用户电量电价弹性矩阵计算的方法, 并通过 IEEE-RTS79 测试系统求得其值, 具有一定的代表性, 因此采用其结果:

$$E = \begin{bmatrix} -0.1732 & 0.1333 & 0.1066 \\ 0.1495 & -0.1150 & 0.0923 \\ 0.2045 & 0.1650 & -0.1320 \end{bmatrix}.$$

实行峰谷分时电价以后, 用户根据不同时段电价的差异自动调整自己的生产方式, 从而减少了峰谷差和负荷波动, 因此用户的峰谷差 ν 和线路损耗比 ω 的最大值均为实行峰谷分时电价之前的数值. 于是用户峰谷差 ν 的最大值为 $\nu_{\max} = 245$ MW, 线路损耗比 ω 的最大值为 $\omega_{\max} = 100\%$. 实行峰谷分时电价以后分别求出二者的最小值: 峰谷差 ν 的最小值 $\nu_{\min} = 138.037$ MW, 线路损耗比 ω 的最小值为 $\omega_{\min} = 97.16\%$. 于是通过求解以式(21)为目标函数, 式(15)~(18)为约束条件的单目标优化模型可以得到如下结果: 当峰电价为 0.934 3 元/(kW·h)、平电价为 0.543 5 元/(kW·h)、谷电价为 0.316 7 元/(kW·h) 时, 综合目标函数 f 取得最小值 0.041 6, 此时峰谷差 ν 为 141.687 MW, 线路损耗比 ω 为 97.23%. 两个目标函数归一后的数值分别为 $\bar{\nu} = 0.0341$ 、 $\bar{\omega} = 0.0247$, 说明通过理想点法求解该优化模型能很好地兼顾到这两个目标. 峰谷分时电价实施前、后负荷对比如图 1 所示.

实行峰谷分时电价以后负荷的峰谷差减少 103.313 MW, 线路损耗减少了 2.77%. 因此, 既减少了线路负荷波动, 提高了电网的运行效率; 也减少了线路损耗, 提高了电网的经济效益. 实施峰谷分时电价后用户的平均用电价格为 0.598 3 元/(kW·h), 用户一天的总负荷为 7 818.8 MW, 和实施之前相比, 用户的平均用电价格略有下降, 用户的总负荷基本保持不变, 说明实施峰谷分时电

价后用户经济上得到了一定的实惠并且用电需求也没受到影响. 实施峰谷分时电价前、后供电公司一天的收入分别为 1 582 266 元、1 550 621 元, 供电公司的收入略有减少, 考虑到实施峰谷分时电价后提高了系统的负荷率, 节省了运行成本, 减少了线路的损耗, 因此供电公司收入的略微减少也是可以接受的.

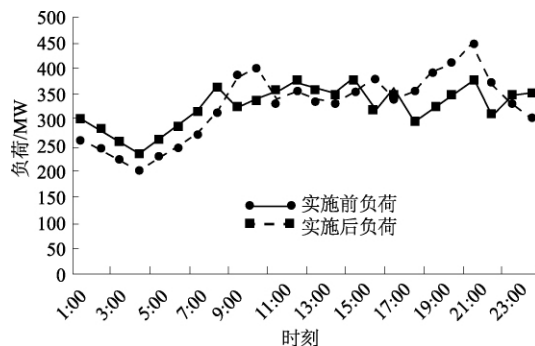


图 1 峰谷分时电价实施前后负荷曲线对比

Fig. 1 Comparison of load curves between before and after the implementation of time-of-use price

5 结论

峰谷分时电价能否顺利实施的关键在于峰平谷时段电价是否合理. 笔者证明了用户总负荷保持一定的基础上, 负荷波动越小线路损耗越小, 然后基于用户电量电价弹性矩阵建立计及线路损耗的峰谷分时电价优化模型. 该优化模型是多目标优化问题, 一般没有绝对最优解, 采用理想点法将其转为单目标模型后求解. 通过实例仿真证明, 计及线路损耗的峰谷分时电价模型能很好地实现削峰填谷、减少线路损耗的目的, 提高了电网运行的效率和经济性; 同时, 由于用户和供电公司的利益均得到了保证, 用户的用电需求也没受到影响, 因此, 优化模型得到的峰平谷分时电价具有一定的合理性, 能够得到顺利的实施.

参考文献:

- [1] ALBADI M H, EI-SAADANY E F. A summary of demand response in electricity markets [J]. Electric power systems research, 2008, 78(11): 1989–1996.
- [2] 李春燕, 许中, 马智远. 计及用户需求响应的分时电价优化模型 [J]. 电力系统及其自动化学报, 2015, 27(3): 11–16.
- [3] 谈金晶, 王蓓蓓, 李扬. 基于多智能体的用户分时电价响应模型 [J]. 电网技术, 2012, 36(2): 257–263.
- [4] 谭忠富, 王绵斌, 乞建勋, 等. 峰谷分时电价优化模型及其模糊求解方法 [J]. 系统工程理论与实践, 2008, 9(9): 145–151.

- [5] 郭联哲,谭忠富,李晓军. 基于用户响应下的分时电价优化设计模型与方法[J]. 电网技术, 2006, 30(5): 25-28.
- [6] 路郑. 国网大刀阔斧节能减排[N/OL]. 中国能源报, 2015-06-22(18).
- [7] 胡福年,汤玉东,邹云. 需求侧实行峰谷分时电价策略的影响分析[J]. 电工技术学报, 2007, 22(4): 168-174.
- [8] 杨桂元,郑亚豪. 多目标决策问题及其求解方法研究[J]. 数学的实践与认识, 2012, 42(2): 108-115.
- [9] 邵璘,周国祥,石雷. 峰谷分时电价决策的优化模型[J]. 统计与决策, 2010(3): 51-53.
- [10] 陈纛,殷善锋. 基于用户需求弹性的峰谷分时电价决策模型[J]. 特区经济, 2013(10): 188-191.
- [11] 陈沧杨,胡博,谢开贵,等. 计入电力系统可靠性与购电风险的峰谷分时电价模型[J]. 电网技术, 2014, 38(8): 2141-2148.

Time-of-use Price Optimization Model Considering Line Loss

ZHAO Guosheng, ZHAN Tianle, LI Bo

(School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The current TOU price optimization models failed to consider the significance of reducing line loss. Through the relationship between line loss and load fluctuation, the targets of the proposed new TOU price optimization model, which was based on price elasticity matrix of demand, were to minimize peak-valley difference and line loss. The optimization model, which was a non-linear multiple objects optimal model, was solved by ideal point method. It was proved by example that the new TOU price optimization model could avoid the peak load and reduce the line loss.

Key words: time-of-use price; line loss; price elasticity matrix of demand; multi-objective optimization model; ideal point method

(上接第32页)

Network Intrusion Detection Algorithm Based on Hybrid Differential Evolution Algorithm

WANG Yaoguang¹, CHEN Weiquan¹, WU Zhenbang¹, QIN Yong², HUANG Han³

(1. Guangdong Dongguan Quality Supervision Testing Center, Dongguan 523000, China; 2. School of Computer Science and Network Security, Dongguan University of Technology, Dongguan 523000, China; 3. School of Software Engineering, South China University of Technology, Guangzhou 510006, China)

Abstract: Intrusion detection algorithm based on machine learning method is one of research hotspot in the field of network equipment testing. In the face of the real-world application requirement, machine learning methods should be further optimized to achieve accurate and stable detection effect. The study optimize steadily several key parameters of least squares support vector machine (SVM) by designing a hybrid differential evolution algorithm with disturbance vector and improved the intrusion detection accuracy and stability of least squares support vector machine (SVM) algorithm by means of adaptive parameter tuning. The experimental results in KDD Cup 09 test set showed that, the proposed network intrusion detection algorithm based on hybrid differential evolution algorithm had better performance on average than many similar algorithm at present.

Key words: network intrusion detection; stability test; hybrid differential evolution; least squares support vector machine

基于 Möbius 变换的复有理圆弧样条

任两品, 薛均晓, 张朝阳, 王定标

(郑州大学 软件与应用科技学院, 河南 郑州 450002)

摘 要: 针对工业产品和机械零件设计中的二次圆弧表示问题, 基于 Möbius 变换的保交比性质和保圆性质, 提出了一种用参数复有理函数精确表示圆弧曲线的方法. 该方法通过构造直线段与圆弧曲线之间的 Möbius 变换, 将圆弧曲线表示为复有理函数的形式, 并结合光滑约束条件构造了复有理圆弧样条函数. 与 Bezier 曲线和 NURBS 曲线等方法相比, 该方法能够精确定义二次圆弧曲线, 而且不需要反求控制顶点和权因子, 实验结果表明所提方法简单易行.

关键词: 计算机数控; 圆弧样条; Möbius 变换

中图分类号: TP391.72

文献标志码: A

doi: 10.3969/j.issn.1671-6833.2017.03.004

0 引言

在工业产品和机械零件的计算机数控加工中, 常常需要使用构造方法简单, 而且容易实现的基本曲线来代替和逼近任意曲线. 由于直线和圆弧的定义形式简单, 而且易于在计算机中进行构造, 所以大多数数控加工中利用直线和圆弧插补的方式来对机械零件进行加工^[1-3]. 特别地, 相对于直线, 圆弧具备更好的光滑性质, 因此使用圆弧曲线来逼近计算机数控的刀具加工路径在近几年得到了大量的研究^[4-7]. 以 Bezier 曲线和样条曲线方法为代表的自由型曲线曲面设计方法, 是图形图像造型的基本工具, 是计算机辅助设计和计算机图形学中运用得最多的基本方法之一. 然而, 该方法在逼近圆弧曲线时计算量巨大, 而且算法的实现过程非常复杂. 例如, 对于具有特定半径的半圆弧, 需要使用 5 次 Bezier 曲线进行逼近, 而且必须反求 Bezier 曲线的控制顶点. 如果圆弧的半径发生改变, 或者对于逼近精度有更高的要求, 则需要重新设定 Bezier 曲线的次数, 而且需要重新计算控制顶点.

NURBS 方法^[8]的出现, 很好地解决了这一问题, 它在保留了样条方法描述自由型形状的长处, 同时, 扩充了样条方法统一表示二次曲线的能力, 包括精确表示二次圆弧. 但是和任何一种方法

一样, NURBS 方法也面临很多问题, 尤其是权因子怎样影响曲线的参数化和怎样确定合适的权因子, 一直是困扰设计工作者并且没有得到解决的问题. 同时, NURBS 方法在定义圆弧的时候需要额外的存储. 例如, 用一个外切正方形作为控制多边形定义一个整圆, 至少需要 7 个控制顶点和 10 个节点. 而传统的表示只要求给出圆心, 半径和垂直于圆所在平面的法矢.

基于三角多项式的 C-curves^[9]解决了圆弧曲线的精确表示问题, 但该方法本身也面临着参数的选取困难问题. 同时, 利用三角多项式表示形式代替有理表示形式, 这将大大增加计算机的运算量. Möbius 变换^[10]最早是由德国几何学家与天文学家莫比乌斯提出来的. 它是复分析中形式结构简单, 具有良好性质, 而且应用广泛的一类保形变换. Möbius 变换具有保交比性质和保圆性质. 在本文中, 通过构造直线段与圆弧曲线之间的 Möbius 变换, 将圆弧曲线表示为复有理函数的形式. 该表示方法形式简单, 存储数据量小, 具有良好的插值性质和仿射不变性质, 而且可以在光滑约束条件下与 B 样条曲线实现连续拼接.

1 Möbius 变换

定义 1.1 由复变函数

收稿日期: 2016-10-30; 修订日期: 2016-12-19

基金项目: 河南省自然科学基金资助项目(162300410262); 河南省高校科技创新团队项目(16IRTSTHN027); 河南省高等学校重点科研项目(18A4130)

通信作者: 薛均晓(1982—), 男, 河南南阳人, 郑州大学副教授, 博士, 主要从事计算机辅助设计、智能制造系统方面的研究, E-mail: xuejx@zzu.edu.cn.

$$w = f(z) = \frac{az+b}{cz+d} \quad (ad-bc \neq 0) \quad (1)$$

所确定的变换称为 Möbius 变换. 其中, a, b, c, d 为复平面上的常数; z 是复变量.

条件 $ad-bc \neq 0$ 是必需的, 否则 w 将蜕化成为常数.

说明: 对 (1) 式进行如下补充定义:

a) 若 $c \neq 0$, 定义 $w = \infty \left(z = -\frac{d}{c} \right)$, 定义

$$w = \frac{a}{c} (z = \infty);$$

b) 若 $c = 0$, 定义 $w = \infty (z = \infty)$.

基于以上补充定义, 我们不难发现式 (1) 能够将扩充的初始平面——地映射为另一个扩充平面, 即由式 (1) 所定义的变换 $w = f(z)$ 能够确定初始扩充 z 平面到目标扩充 w 平面之间的一一对应变换关系.

定理 1.1 设 Möbius 变换将扩充 z 平面 3 个不同的点 z_1, z_2, z_3 对应变换成扩充 w 平面上的 3 个点 w_1, w_2, w_3 , 即 $w_j = f(z_j), j = 1, 2, 3$. 则该 Möbius 变换就被唯一确定, 并且可以写成下述形式:

$$w = f(z) = \frac{[(z-z_3)(z_1-z_2)w_2(w_1-w_3) - (z-z_2)(z_1-z_3)w_3(w_1-w_2)]}{[(z-z_3)(z_1-z_2)(w_1-w_3) - (z-z_2)(z_1-z_3)(w_1-w_2)]}. \quad (2)$$

由定理 1.1 可得, 扩充平面上的 3 对对应点能够唯一确定一个 Möbius 变换.

2 基于 Möbius 变换的圆弧曲线表示方法

定理 2.1 设 $v_1(x_1, y_1), v_2(x_2, y_2), v_3(x_3, y_3)$ 是平面上不在同一条直线上的任意 3 点. 记 $w_j = x_j + iy_j (j=1, 2, 3)$, w_j 是点 v_j 的复数表示, 则下述复有理函数

$$w(t) = \frac{[(t-t_3)(t_1-t_2)w_2(w_1-w_3) - (t-t_2)(t_1-t_3)w_3(w_1-w_2)]}{[(t-t_3)(t_1-t_2)(w_1-w_3) - (t-t_2)(t_1-t_3)(w_1-w_2)]}, t \in [t_1, t_3], \quad (3)$$

精确表示由 v_1, v_2, v_3 这 3 点所确定的一段圆弧曲线, 并且满足 $w_j = w(t_j) (j=1, 2, 3)$. 其中, $t_j \in \mathbb{R} (j=1, 2, 3)$ 为实数节点, 且 $t_1 < t_2 < t_3$.

证明 因为 $t \in [t_1, t_3] \subset \mathbb{R}$, 不妨将 t 看作虚部为零的复数, 则式 (3) 所确定的为 Möbius 变换.

由 Möbius 变换的保交比性质, 4 点 w, w_1, w_2, w_3 的交比 (w, w_1, w_2, w_3) 等于 4 点 t, t_1, t_2, t_3 的交比 (t, t_1, t_2, t_3) . 显然, (t, t_1, t_2, t_3) 为实数, 所以 (w, w_1, w_2, w_3) 也为实数. 结合 Möbius 变换的保圆周性质, 4 点 w, w_1, w_2, w_3 位于同一圆周上, 即

式 (3) 所定义的函数精确表示由 v_1, v_2, v_3 这 3 点所确定的一段圆弧曲线.

事实上, 在定理 2.1 中, 若记经过 v_1, v_2, v_3 3 点的圆弧曲线为 Γ , 则式 (3) 建立了实数轴上的区间 $[t_1, t_3]$ 与圆弧曲线 Γ 之间的一一对应关系. 即式 (3) 为圆弧曲线 Γ 的参数复有理函数表示. 换句话说, 任给平面上不在同一直线上的 3 点, 都可以利用式 (3) 来构造由它们所确定的圆弧曲线.

下面说明式 (3) 是非退化的. 式 (3) 可以整理为 M-形式, 定义如下:

$$w(t) = \frac{at+b}{ct+d},$$

式中,

$$\begin{aligned} a &= (t_1 - t_2)w_2(w_1 - w_3) - (t_1 - t_3)w_3(w_1 - w_2); \\ b &= t_2(t_1 - t_3)w_3(w_1 - w_2) - t_3(t_1 - t_2)w_2(w_1 - w_3); \\ c &= (t_1 - t_2)(w_1 - w_3) - (t_1 - t_3)(w_1 - w_2); \\ d &= t_2(t_1 - t_3)(w_1 - w_2) - t_3(t_1 - t_2)(w_1 - w_3); \\ t_1 &< t_2 < t_3, t \in [t_1, t_3]. \end{aligned}$$

由于

$$ad - bc = (t_1 - t_2)(w_1 - w_2)(t_2 - t_3)(w_2 - w_3)(t_1 - t_3)(w_1 - w_3) \neq 0,$$

所以 M-形式是非退化的.

记式 (3) 为参数复有理圆弧曲线, 下文中将讨论它的相关性质.

3 参数复有理圆弧曲线的性质

3.1 仿射不变性

设

$$\begin{cases} \Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \\ w \rightarrow w \cdot M + N, \end{cases} \quad (4)$$

式中, M 为 2×2 矩阵; N 为 1×2 向量.

把 $\Phi(w_j) = w_j \cdot M + N$ 代入式 (4), 经过简单计算可得

$$w(t; w_1 \cdot M + N, w_2 \cdot M + N, w_3 \cdot M + N) = w(t; w_1, w_2, w_3) \cdot M + N.$$

亦即

$$w[t; \Phi(w_1), \Phi(w_2), \Phi(w_3)] = \Phi[w(t; w_1, w_2, w_3)]. \quad (5)$$

3.2 几何不变性

设

$$\begin{cases} \varphi: \mathbb{R} \rightarrow \mathbb{R}; \\ t \rightarrow kt + h, \end{cases} \quad (6)$$

其中, k, h 均为实数.

把 $\varphi(t) = kt + h, \varphi(t_j) = kt_j + h, (j=1, 2, 3)$ 代入式 (5), 经过简单计算可得

$$w[\varphi(t); \varphi(t_1), \varphi(t_2), \varphi(t_3)] = w(t). \quad (7)$$

上述性质表明: M-形式中的参数定义区间和节点不是本质的, 可以根据需要选择合适的参数定义区间以及节点. 为了方便, 不妨设

$$t_1 = 0; t_2 = \frac{1}{2}; t_3 = 1,$$

参数定义区间为 $[0, 1]$. 则圆弧曲线可以表示为

$$w(t) = \frac{(t-1)w_2(w_1-w_3) - (2t-1)w_3(w_1-w_2)}{(t-1)(w_1-w_3) - (2t-1)(w_1-w_2)}, t \in [0, 1]. \quad (8)$$

3.3 微分性质

经过简单计算, 可得

$$\frac{dw}{dt} = \frac{ad-bc}{(ct+d)^2} =$$

$$\frac{(t_1-t_2)(w_1-w_2)(t_2-t_3)(w_2-w_3)(t_1-t_3)(w_1-w_3)}{[(t-t_3)(t_1-t_2)(w_1-w_3) - (t-t_2)(t_1-t_3)(w_1-w_2)]^2},$$

所以

$$\left. \frac{dw}{dt} \right|_{t=t_1} = \frac{t_2-t_3}{(t_1-t_3)(t_1-t_2)} \cdot \frac{(w_1-w_2)(w_1-w_3)}{w_2-w_3};$$

$$\left. \frac{dw}{dt} \right|_{t=t_3} = \frac{t_1-t_2}{(t_1-t_3)(t_2-t_3)} \cdot \frac{(w_1-w_3)(w_2-w_3)}{w_1-w_2},$$

如果 $t_1 = 0, t_2 = \frac{1}{2}, t_3 = 1$, 则

$$w'(0) = -\frac{(w_1-w_2)(w_1-w_3)}{w_2-w_3};$$

$$w'(1) = -\frac{(w_1-w_3)(w_2-w_3)}{w_1-w_2}.$$

4 参数复有理圆弧样条

4.1 构造 C^0 连续的圆弧样条

已知平面上 n 个点 $v_1(x_1, y_1), v_2(x_2, y_2), \dots, v_n(x_n, y_n), n \geq 3$, 将其表示为复数形式 $w_1 = x_1 + iy_1, w_2 = x_2 + iy_2, \dots, w_n = x_n + iy_n$. 定义如下:

$$l_k(t; w_j, w_{j+1}, w_{j+2}) = \frac{(1-t)t_k^* w_{j+1}(w_j - w_{j+2}) + (t-t_k^*) w_{j+2}(w_j - w_{j+1})}{(1-t)t_k^* (w_j - w_{j+2}) + (t-t_k^*) (w_j - w_{j+1})}, \quad (9)$$

其中, $t \in [0, 1]$ 是参数; $t_k^* = \frac{l_j v_{j+1}}{l_j v_{j+2}}$ 是区间 $[0, 1]$

内的节点; 下标 $j = 1, 2, \dots, n-2; k = 1, 2, \dots, n_1$.

计算可得 $n = 2n_1 + 1$. 根据定理 2.1, 显然式 (9) 定义了一条插值于已知 n 个点的参数复有理圆弧样条曲线, 而且由参数复有理圆弧曲线的微分性质不难验证该曲线是 C^0 连续的.

在本文中, 笔者构造了两条 C^0 连续的参数复有理圆弧样条曲线, 如图 1 所示. 在图 1 中, 样条曲线在相邻的连接点处满足 C^0 连续的连接条件, 而且精确插值于给定的型值点. 显然, 如果

使用线性插值方法, 对于给定的 7 个型值点, 必须构造 6 条直线段, 而基于 3 点插值的复有理圆弧样条曲线仅仅只需要构造 3 段圆弧曲线, 而且在除了拼接点处满足 C^0 连续的连接条件外, 圆弧样条曲线在其他点处是 C^∞ 连续的.

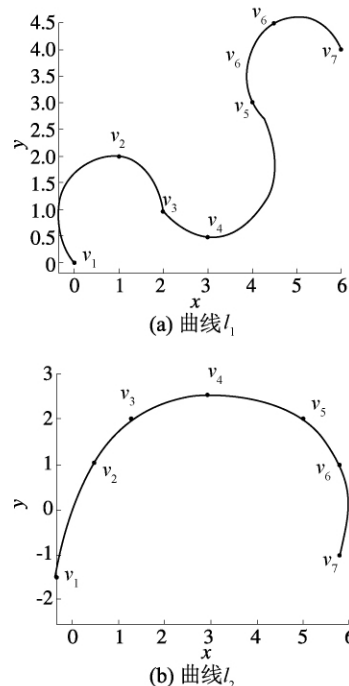


图 1 C^0 连续的圆弧样条

Fig. 1 C^0 continuous circular arc representation

4.2 构造 GC^1 连续的圆弧样条

已知平面上 5 个点 $v_j(x_j, y_j) (j = 1, 2, \dots, 5)$, 将其表示为复数形式: $w_j = x_j + iy_j (j = 1, 2, \dots, 5)$, 构造两段圆弧曲线:

$$l_1(t; w_1, w_2, w_3) = \frac{(1-t)t_1^* w_2(w_1-w_3) + (t-t_1^*) w_3(w_1-w_2)}{(1-t)t_1^* (w_1-w_2) + (t-t_1^*) (w_1-w_3)} (t \in [0, 1]);$$

$$l_2(t; w_3, w_4, w_5) = \frac{(1-t)t_2^* w_4(w_3-w_5) + (t-t_2^*) w_5(w_3-w_4)}{(1-t)t_2^* (w_3-w_5) + (t-t_2^*) (w_3-w_4)} (t \in [0, 1]).$$

其中, $t_1^* = \frac{l_1 v_2}{l_1 v_3}, t_2^* = \frac{l_2 v_4}{l_2 v_5}$. 圆弧曲线 $l_1(t; w_1, w_2, w_3)$ 经过点 v_1, v_2, v_3 , 圆弧曲线 $l_2(t; w_3, w_4, w_5)$ 经过点 v_3, v_4, v_5 . 当且仅当下述条件:

$$\frac{(w_2-w_3)(w_3-w_1)}{(w_1-w_2)} = \alpha \cdot \frac{(w_3-w_4)(w_3-w_5)}{(w_4-w_5)}$$

成立时, 两段圆弧曲线 l_1 和 l_2 在连接点 v_3 处是 GC^1 连续的, 其中, $\alpha \in \mathbb{R}$ 为一任意常数.

已知平面上 n 个点 $v_1(x_1, y_1), v_2(x_2, y_2), \dots, v_n(x_n, y_n), (n \geq 3)$, 将其表示为复数形式 $w_1 = x_1 + iy_1, w_2 = x_2 + iy_2, \dots, w_n = x_n + iy_n$. 类似于前

文中 C^0C^0 连续的圆弧样条曲线的构造,首先定义下述圆弧曲线.

$$l_k(t; w_j, w_{j+1}, w_{j+2}) = \frac{(1-t)t_k^* w_{j+1}(w_j - w_{j+2}) + (t-t_k^*) w_{j+2}(w_j - w_{j+1})}{(1-t)t_k^* (w_j - w_{j+2}) + (t-t_k^*) (w_j - w_{j+1})}, \quad (10)$$

式中: $t \in [0, 1]$ 是参数; $t_k^* = \frac{l_j v_j v_{j+1}}{l_j v_j v_{j+2}}$ 是区间 $[0, 1]$ 内的节点; 下标 $j = 1, 2, \dots, n-2$; $k = 1, 2, \dots, n_1$.

若给定的 n 个点满足下述条件:

$$\frac{(w_{j+1} - w_{j+2})(w_{j+2} - w_j)}{(w_j - w_{j+1})} = \alpha \cdot \frac{(w_{j+2} - w_{j+3})(w_{j+3} - w_{j+4})}{(w_{j+3} - w_{j+4})}, \quad (11)$$

式中, $\alpha \in \mathbb{R}$ 为一任意常数; $j = 1, 2, \dots, n-4$. 则由式(10)定义的圆弧样条曲线是 GC^1 连续的,且过已知的 n 个点.

图2为两条 GC^1 连续的圆弧样条曲线 l_1 和 l_2 . 其中,圆弧样条曲线 l_1 由3段圆弧曲线 l_{11} 、 l_{12} 和 l_{13} 组成,对于平面上给定的7个型值点 $v_1(-3.5, -1.5)$ 、 $v_2(-2.5, 1)$ 、 $v_3(-1.5, 2)$ 、 $v_4(0.5, 2.5)$ 、 $v_5(2.6852, 0.3484)$ 、 $v_6(2.1156, -1.4266)$ 、 $v_7(1.7729, -1.6972)$, 圆弧曲线 l_{11} 插值于 v_1 、 v_2 、 v_3 , 圆弧曲线 l_{12} 插值于 v_3 、 v_4 、 v_5 , 圆弧曲线 l_{13} 插值于 v_5 、 v_6 、 v_7 . 经简单计算可得,圆弧曲线 l_{11} 和 l_{12} 在

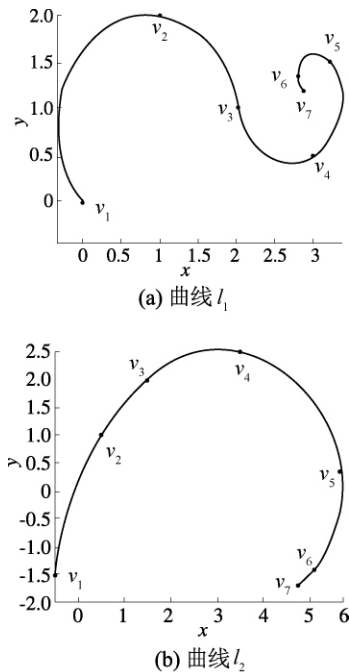


图2 GC^1 连续的圆弧样条

Fig. 2 GC^1 continuous circular arc representation

点 v_3 处为 GC^1 连续,圆弧曲线 l_{12} 和 l_{13} 在点 v_5 处为 GC^1 连续. 如果使用NURBS曲线构造 GC^1 连续的圆弧样条曲线 l_1 ,则需要使用2次NURBS曲线,还必须利用反求矩阵计算NURBS曲线的控制顶点和权因子,需要大量的额外运算;而且权因子选择的不同,会影响插值样条曲线的唯一性. 但本文方法只需要把各个型值点转换为复数形式,直接代入样条曲线方程函数即可,简单易行.

5 结束语

笔者利用Möbius变换的保交比性质和保圆性质,提出了一种精确定义圆弧曲线的参数复有理函数方法,并基于该方法,研究了具有 C^0 连续和 GC^1 连续性质的参数复有理圆弧样条曲线的构造方法. 与目前常用的NURBS曲线和C-curves曲线方法相比,笔者提出的圆弧曲线表示方法不需要权因子和额外的控制参数,故而简便易行.

参考文献:

- [1] 周君,任景刚,廖铃吉,等. 航空发动机壳体2.5轴铣削加工刀具优选方法[J]. 航空制造技术,2016(1): 143-146.
- [2] 陈银清,郑泽钿. 基于MasterCAM自由曲面加工刀具路径优化[J]. 现代制造工程,2014(5): 46-51.
- [3] 吕勇刚,汪国昭,杨勋年,等. 单圆弧样条保形插值算法[J]. 计算机学报,2002,25(5): 483-486.
- [4] AHLFORS L V. Complex analysis [M]. New York, USA: McGraw-Hill Book Co, 1979.
- [5] 施法中. 计算机辅助几何设计与非均匀有理B样条[M]. 北京: 北京航空航天大学出版社, 1994.
- [6] 朱晓临. 一种参数有理圆弧样条[J]. 合肥工业大学学报(自然科学版), 2003, 26(2): 180-184.
- [7] YEUNGAND M, WALTON D J. Curve fitting with arc-splines for NC tool path generation [J]. Computer-aided design, 1994, 26(11): 845-849.
- [8] 王琦,郭非,王启义. 圆弧样条逼近为机械零件几何轮廓的自动编程[J]. 机械工程学报, 1998, 34(2): 20-25.
- [9] LI Z, MEEK D S. Smoothing an arc spline [J]. Computers & graphics, 2005, 29(4): 576-587.
- [10] 齐倩,方美娥. G_2 连续的圆弧样条曲线插值[J]. 杭州电子科技大学学报, 2012(6): 45-48.

(下转第58页)

文章编号: 1671-6833(2017)06-0054-05

基于 FPGA 的直流电机伺服控制系统设计与实现

常怡萍, 刘 琴, 杨 蕾

(中原工学院 电子信息学院 河南 郑州 450007)

摘 要: 为了满足直流电机伺服控制系统实时控制的要求,提高电机跟踪控制的精度,设计了基于单片 FPGA 的直流电机伺服控制系统. 该控制系统采用 Nios II 内核实现位置环和速度环控制策略,系统的电流环通过并行硬件电路实现. 为了减小电机的力矩波动,设计了基于 Anti-Windup 策略的 PI 速度控制器;为了减小系统的超调量和调节时间,位置环采用了超前滞后校正器. 仿真和实验结果表明,设计的伺服控制系统具有较好的稳态精度和动态性能,满足系统的设计要求.

关键词: FPGA; Anti-Windup; Biss; 伺服控制

中图分类号: TP 273

文献标志码: A

doi: 10.3969/j.issn.1671-6833.2017.03.006

0 引言

直流电机以其优异的性能在航空航天、数控机床、汽车电子等领域得到了广泛的应用. 传统的直流电机伺服控制系统通常采用基于 DSP 和 FPGA 架构的控制器,此类控制器电路结构复杂、灵活性较差,不利于系统的升级和维护. 为了简化外围硬件电路设计,近年来基于单片 FPGA 的伺服控制器成为研究的热点,并且在工业控制领域得到了成功的应用^[1-3]. 文献[4]采用 FPGA 完成了基于 Anti-Windup 策略的永磁同步电机速度控制,实现了永磁同步力矩电机转速的快速无超调控制;文献[5]在 FPGA 中实现了感应电机的神经网络控制,实现了感应电机的集成伺服控制;文献[6]通过硬件语言在 FPGA 中完成了永磁同步直线电机的自适应模糊控制^[6]. 上述控制算法的实现均采用并行硬件电路实现,调试周期较长,对编程能力要求较强. 由于 FPGA 支持基于 C 语言的 Nios II 内核,该内核具有快速的运算性能和灵活的外围接口设计,因此,Nios II 内核在伺服控制系统中得到了较多的应用^[7].

传统的 PI 控制器受控制器输出限幅的影响,电机的力矩和转速输出容易出现超调和振荡,文献[8]采用基于 Anti-Windup 策略的 PI 控制器,实现了望远镜伺服控制系统的快速、无超调控制. 文献[9]通过新型的 Anti-Windup 策略

完成控制器从 P 到 PI 的平稳切换,实现了大阶跃信号的无超调控制.

为了提高直流电机伺服控制系统的集成度和系统的实时性,笔者将采用 Altera 公司的 FPGA-EP4C22E22 对直流电机伺服控制系统进行设计,在 FPGA 中实现了速度环和位置环控制策略、电流环策略、电机电流采集、脉宽调制(PWM)、Biss 编码器数据采集. 相比于传统基于 DSP 的伺服控制系统,提出的方案在精简硬件电路设计的同时,具有更好的灵活性. 针对速度环阶跃响应超调的问题,采用了具有 Anti-Windup 策略的 PI 控制器,最后通过仿真和实验对上述设计进行了验证.

1 伺服系统的硬件设计方案

直流电机伺服控制系统的总体硬件设计方案如图 1 所示,其中电机的位置环和速度环在 Nios II 内核中实现,电流环 PI 控制器采用并行电路实现,PWM 由比较器产生,电机电流数据和 Biss 编码器数据的采集通过有限状态机实现. 在 FPGA 中采用有限状态机可以更有效地按照芯片 AD7656 数据手册要求的时序进行 AD 数据的采样和读取. 上述模块的时钟统一由倍频/分频模块产生. 硬件设计的难点是电流 PI 控制器、电流 AD 采集的有限状态机实现和 Biss 编码器数据采集,因此,下面重点介绍这两部分的设计.

收稿日期: 2016-05-10; 修订日期: 2016-06-28

基金项目: 国家自然科学基金资助项目(61501530)

作者简介: 常怡萍(1979—),女,河南周口人,中原工学院讲师,主要从事扩频通信、宽带无线通信研究, E-mail: chanyiping@126.com.

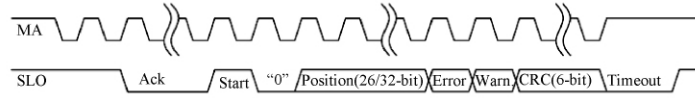


图4 32位绝对式编码器数据格式

Fig. 4 Data mode of 32 bits absolute encoder

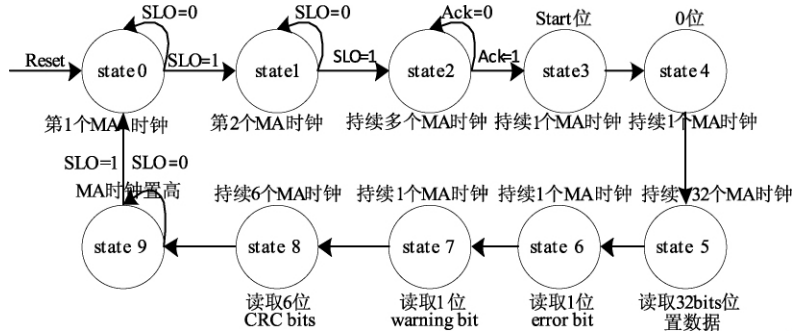


图5 32位编码器读取状态机

Fig. 5 FSM of 32 bits absolute encoder data reading

图中 Ack 为应答信号,其持续时间为读数头计算绝对位置的时间。“start”位和“0”位均为1位,“start”位始终是高电平,“0”位始终是低电平;绝对位置数据为二进制格式,首先发送给最高位.对于旋转编码器来说,每转正好有 $2n$ 个脉冲,之后脉冲数溢出绕回到零。“error”位低电平有效。“1”表示传输的位置信息已被读数头的内置安全校验算法校验,结果正确。“0”表示内部检查失败,位置信息不可信。“warning”位低电平有效:“0”表示应对编码器(及/或读数窗口)进行清洁。“CRC”位为校验位,位置、错误及警告数据的CRC多项式为: $x^6 + x^1 + x^0$,它先被传输为最高位,然后转换.起始位和“0”位从CRC计算中忽略.

2 位置和速度控制器设计

伺服控制系统采用如图6所示的控制结构,由位置控制器、速度控制器、结构滤波器、速度检测和控制模型组成.其中控制模型包括了电流环和系统的结构模态.为了增加系统的带宽和响应速度,需要设计位置控制器、速度控制器和结构滤波器,位置控制器采用超前滞后补偿的形式,速度控制器采用具有 Anti-Windup 策略的 PI 控制器.结构滤波器的应用目的是抑制在谐振频率处的驱动力矩成分,提高控制系统的增益,进而提高系统的带宽和动态性能.结构滤波器的表达式为

$$W_{SF}(s) = \frac{s^2 + 2\zeta_Z \omega_{NF} s + \omega_{NF}^2}{s^2 + 2\zeta_P \omega_{NF} s + \omega_{NF}^2}, \quad (4)$$

式中: ω_{NF} 为弹簧质量模型的谐振频率,称为陷波频率; ζ_Z 、 ζ_P 分别为结构滤波器的零点和极点阻尼

系数.

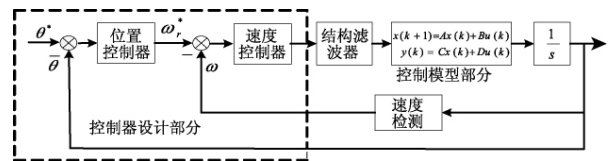


图6 直流电机伺服控制系统结构图

Fig. 6 Dc motor servo control system structure

位置超前滞后补偿控制器形式为

$$G_{pc}(s) = \frac{k_p (T_1 s + 1) (T_3 s + 1)}{(T_2 s + 1) (T_4 s + 1)}, \quad (5)$$

式中: k_p 为控制器增益系数; T_1 、 T_2 、 T_3 、 T_4 为控制器时间常数.具有 Anti-Windup 策略的速度 PI 控制器的输出 i_q 为

$$i_q = k_p \cdot \Delta\omega + q, \quad (6)$$

式中: q 表示 PI 控制器的积分输出; $\Delta\omega = \omega^* - \omega$. 控制器输出工作在线性区时,速度环采用 PI 控制形式;当速度阶跃信号较大时,PI 控制器输出会达到饱和,基于反馈计算的积分输出值将以一定的速率复位为零,此时,控制器为纯比例控制.

积分输出 q 的表达式为

$$q = k_p \cdot \Delta\omega + \frac{k_i}{s} (\Delta\omega - k_a \cdot \Delta i_q), \quad (7)$$

式中: $\Delta i_q = i_q - i_q^*$; k_p 、 k_i 分别为 PI 控制器的比例时间常数和积分时间常数; k_a 为 PI 控制器反馈计算常数.控制算法由电流环、速度环和位置环组成.电流环需要实时、快速、准确地跟踪电流指令信号,因此电流控制器采用 PI 控制器,以使系统快速响应,消除静态误差.速度环的作用是增强系统的抗负载扰动的能力,抑制速度波动,实现稳态

无静差.位置环的作用是实现电机角位置的精确定位,为了保证系统的稳定,位置控制器采用超前-滞后补偿控制器.

3 实验结果分析

基于上述设计对直流电机驱动的小型转台进行了伺服控制实验,电机参数如下:电阻 $R = 1.6 \Omega$,电感 $L = 6.0 \text{ mH}$,负载转动惯量为 $J = 13 \text{ kg} \cdot \text{m}^2$;编码器采用 32 位 Biss 绝对式编码器,分辨率为 0.0003° .位置环设计参数: $k_p = 150$, $T_1 = 1.5$, $T_2 = 0.039$, $T_3 = 0.29$, $T_4 = 2.9$;速度环设计参数: $k_p = 600$, $k_i = 1500$, $k_a = 0.5$.位置环和速度环采样周期均为 1000 Hz .实验测得伺服系统的速度、位置阶跃响应曲线如图 7、8 所示.

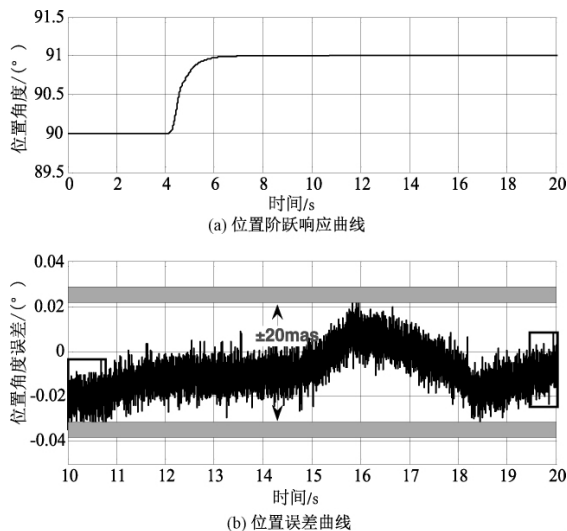


图7 伺服系统的位置阶跃响应曲线

Fig.7 Position step response curve of the servo system

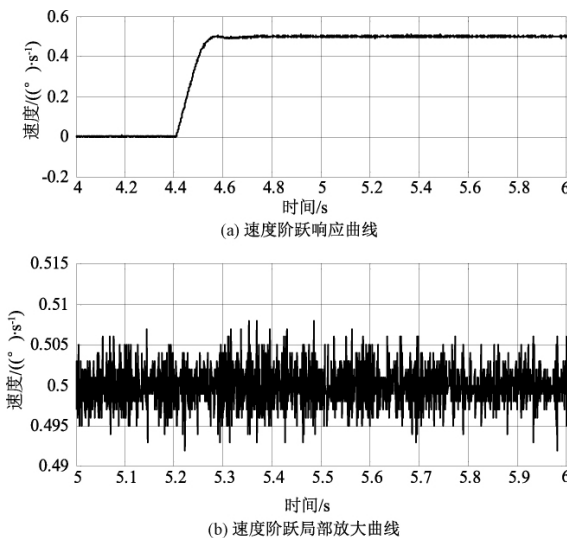


图8 伺服系统的速度阶跃响应曲线

Fig.8 Speed step response curve of the servo system

由图 7 可以看出,位置阶跃具有较快的动态响应能力,位置定位时快速、无超调,稳态时位置误差为 $\pm 0.02^\circ$ (20 毫角秒),满足伺服系统精确定位的要求.伺服系统在 $0.5^\circ/\text{s}$ 时的速度阶跃响应曲线如图 8 所示,速度阶跃响应调节时间 $t_s < 0.2 \text{ s}$,稳态速度波动百分比为 0.54% ,伺服系统的速度阶跃响应曲线说明系统的速度环在采用具有 Anti-Windup 策略的控制器后,具有较好的动态响应能力和较高的稳态精度.

4 结论

设计了基于单片 FPGA 的直流电机伺服控制系统,简化了外围硬件设计,在 Nios II 内核中实现了位置超前滞后控制器;设计了具有 Anti-Windup 策略的速度控制器和结构滤波器,减小了系统的速度超调量和调节时间.最后,通过实验验证了设计的伺服控制系统的良好性能,为直流电机伺服控制系统的集成化设计提供了较好的设计参考.

参考文献:

- [1] ZHANG D, LI H. A stochastic-based FPGA controller for an induction motor drive with integrated neural network algorithms [J]. IEEE Trans Ind Electron, 2008, 55(2): 551-561.
- [2] KUNG Y S, HUANG C C, TSAI M H. FPGA realization of an adaptive fuzzy controller for PMLSM drive [J]. IEEE Trans Ind Electron, 2009, 56(8): 2923-2935.
- [3] JEZERNIK K, KORELIC J, HORVAT R. PMSM sliding mode FPGA-based control for torque ripple reduction [J]. IEEE Trans Power Electron, 2013, 28(7): 3549-3556.
- [4] HANUS R, KINNAERT M, HENROTTE J L. Conditioning technique, a general anti-windup and bumpless transfer method [J]. Automatica, 1987 (23): 729-739.
- [5] JEZERNIK K, KORELI J, HORVAT R. PMSM sliding mode FPGA-based control for torque ripple reduction [J]. IEEE Trans Power Electron, 2013, 28(7): 3549-3556.
- [6] LIN F J, TENG L T, CHANG C K. Adaptive backstepping control for linear-induction-motor drive using FPGA [J]. Electr Power Appl, 2006(153): 484-492.
- [7] KUNG Y S, TSAI M H. FPGA-based speed control IC for PMSM drive with adaptive fuzzy control [J]. IEEE Trans Power Electron, 2007(22): 2476-2486.
- [8] 李洪文. 基于内模 PID 控制的大型望远镜伺服系统 [J]. 光学精密工程, 2009, 17(2): 328-332.

- [9] CHOI J W, LEE S C. Antiwindup strategy for PI-Type speed controller[J]. IEEE Trans Ind Electron, 2009 (56) : 2039 – 2046.
- [10] 霍海龙. 基于 BiSS 协议的编码器及其在伺服驱动系统中的应用[J]. 电气自动化, 2011, 33(3) : 12 – 15.
- [11] 肖鲲,王莉娜. 基于 FPGA 的 BiSS 协议光电编码器通信模块设计[J]. 电子测量技术, 2008, 31(7) : 188 – 191.

Design and Implementation of Motor Servo Control System Based on FPGA

CHANG Yiping, LIU Qin, YANG Lei

(School of Electronic and Information Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China)

Abstract: In order to meet the real-time control requirement of servo system, a dc motor control system based on FPGA was designed. The position and speed controller were realized in the Nios II, and the current loop was executed in FPGA. In order to reduce the torque ripple, an Anti-Windup PI controller was designed. The lead-lag controller was adopted for the position controller. The simulation and experiment result showed that the servo control system had better steady and dynamic performance, which met the requirement of the servo control system.

Key words: FPGA; Anti-Windup; biss; servo control

(上接第 53 页)

Circular Arc Representation Based on Möbius Transformation

REN Liangpin, XUE Junxiao, ZHANG Chaoyang, WANG Dingbiao

(School of Software, Zhengzhou University, Zhengzhou 450002, China)

Abstract: In CNC machinery, the industrial products and machinery parts are often represented by circular arcs. A new method for representing circular arc based on Möbius transformation was presented in the paper. By constructing the Möbius transformation between a straight line segment and a circular curve segment, the arc curve was expressed as a form of complex rational function, and a complex rational arc spline function was also constructed based on the smooth constraints and the Möbius transformation. The representation had no weight factors or control parameters, and it was geometric and affine invariant. Compared with the classical method for representing circular arc, such as NURBS or C-curves, the presented method was much simpler.

Key words: CNC; circular arc; Möbius transformation

文章编号: 1671-6833(2017)06-0059-05

考虑顾客时间紧迫度的生鲜电商配送路径优化问题

崔岩, 张子祥, 时新, 王晓亮, 王振锋

(河南农业大学 机电工程学院, 河南 郑州 450002)

摘要: 针对不确定环境下生鲜电商路径优化问题, 考虑客户决策的有限理性特点, 以累积前景理论为基础, 针对生鲜电商的特点, 构造以货损成本、惩罚成本、运输成本最小为目标, 以代理点需求、车辆装载质量和客户要求服务时间窗为约束, 建立了生鲜电商配送路径优化模型. 采用改进粒子群算法对模型进行求解. 计算结果表明: 基于累积前景理论的模型能够更有效的满足客户对时间窗的要求, 会相应的提高配送服务的满意度并在一定程度上提高企业效益, 为生鲜电商企业优化配送路径提供理论支撑.

关键词: 累积前景理论; 车辆路径问题; 生鲜电商; 时间窗; 改进粒子群算法

中图分类号: TP29

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.008

0 引言

国内生鲜电商的市场发展前景广阔, 但盈利情况却仅有1%^[1]. 影响生鲜电商盈利的因素主要有: 我国冷链物流发展的滞后, 生鲜电商配送成本过高. 而带时间窗的车辆路径问题(VRPTW)分为软时间窗和硬时间窗, 是对VRP问题进一步的改进. 目前, 对车辆路径的研究大多采用启发性算法, 并取得了一定的成果. 但是生鲜电商配送相较于传统配送更看重于配送的及时性^[2]. 生鲜电商的配送路径优化还处于起步阶段. 龚树生, 梁怀兰^[3]等对生鲜食品的冷链物流网络进行了研究.

然而常用的路径优化方案忽略了天气、路况、人为等不确定因素的影响, 导致配送方案结果存在一定偏差. 配送路径优化很大程度上依赖于交通系统, 而交通系统容易受供给和需求两方面影响^[4]. 不确定条件下人们的决策行为容易受到个人习惯偏好、对待风险态度的影响. Kahneman等在有限理性假设的基础上提出累积前景理论(cumulative prospect theory, CPT)^[5]. 目前基于累积前景理论研究路径优化的文献不多, 任亮等^[6]基于累积前景理论建立了考虑运输时间的4PL路径优化模型. 笔者主要针对顾客对时间窗的细化

要求, 建立了一个通过灵活调节惩罚系数来满足不同顾客对时间窗的要求同时节省商家配送成本的模型.

1 问题描述

生鲜电商的配送问题可以描述为, 顾客经网站或手机APP下单后, 商家安排配送. 但不论是商家自行配送或者经过第三方配送, 都是先到相应的代理点(或自营点、自提点), 然后再根据客户地址进行细化配送的, 笔者基于此研究同城配送中心向各个代理点配送的路径优化问题. 相较于经典VRP问题^[7]增添的约束条件如下: ①针对生鲜食品的易腐特点增加了货损成本; ②针对电商配送的时间特性增加了惩罚成本; ③运用累积前景理论确定单位惩罚系数; ④每辆车载重量相同, 行驶速度恒定, 即配送时间仅与配送距离有关.

1.1 模型参数和变量

1.1.1 参数

c_{ij} 为从点*i*到点*j*的运输成本; l 为配送中心的车辆数; b_k 为每辆车的载重量; w_i 为各代理点的需求量; c_s 为货损成本; α 为单位时间货损系数; p 为产品单价; L_{vi} 为客户允许服务开始时刻; U_{vi} 为客户允许服务延迟时刻; $[e_{is}, e_{ie}]$ 为每个代理点*i*带

收稿日期: 2017-04-15; 修订日期: 2017-08-20

基金项目: 国家自然科学基金资助项目(71001111); 河南省科技厅科技攻关国际合作项目(30600802); 2016年度河南农业大学科技创新基金资助项目(KJCX2016A04)

作者简介: 崔岩(1965—), 女, 河南南阳人, 河南农业大学副教授, 主要从事农产品物流、交通运输等方面研究.

通信作者: 王振锋(1982—), 男, 河南淅川人, 河南农业大学副教授、博士, 主要从事农产品物流研究, E-mail: 285291380@qq.com.

有的时间窗; t_{jl} 为车辆 l 从出发点到代理点 j 的时间; θ_{ei}, θ_{li} 为单位惩罚系数; s_{0l} 为车辆 l 出发时刻; s_{il} 为车辆 l 到达代理点 i 的时刻; t_{ij} 为代理点 i 到 j 的运行时间; s_{ei} 为对代理点 i 服务开始的时刻; s_{li} 为完成服务的时刻; t_i 为车辆完成代理点 i 配送任务所需要的服务时间; t_{\max} 为最大等待时间; t_e 表示服务等待时间; Q_i 为惩罚成本, 定义如下:

$$Q_i = \begin{cases} \infty, & s_{ei} \leq L_{vi}; \\ \theta_{ei} t_e, & L_{vi} \leq s_{ei} < e_i; \\ 0, & e_{is} \leq s_{ei} \leq e_{ie}; \\ \theta_{li} t_e, & e_{is} < s_{ei} \leq U_{vi}; \\ \infty, & s_{ei} \geq U_{vi}. \end{cases} \quad (1)$$

1.1.2 决策变量

将配送模型抽象为一个多重图 $R(I, E)$. I ($|I| = n$) 表示配送中心及各个代理点, 其中配送中心编号为 0, 各代理点编号为 $1, 2, \dots, n$; E 表示可选的配送线路的集合. 配送任务及配送中心都以点 i ($i = 0, 1, 2, \dots, n$) 来表示. 定义变量如下:

$$\begin{cases} x_{ijl} = \begin{cases} 1, & \text{车辆 } l \text{ 从点 } i \text{ 行驶到点 } j; \\ 0, & \text{其他,} \end{cases} \\ y_{li} = \begin{cases} 1, & \text{代理点 } i \text{ 的任务由车辆 } l \text{ 完成;} \\ 0, & \text{其他.} \end{cases} \end{cases} \quad (2)$$

电商物流运输时间很难精准定位, 一般以离散时间进行描述. 因此, 将总的配送时间 t_{ijl} 设为一个时间段上的离散型随机变量, 服从 n 个点上的均匀分布, 且相互独立.

G 表示一条从起点到目的节点的通路, 即优化后的配送线路; $t(G)$ 为 G 的运输时间; k 表示 G 所包含的代理点个数 ($k \in n$).

$$t(G) = \sum_i \sum_j \sum_l t_{ijl} x_{ijl} + (k-1) t_i. \quad (3)$$

2 基于 CPT 的决策方案

由 VRPSTW 即软时间窗车辆路径问题可知, 当配送时间介于 L_{vi} 和 e_i , 以及 L_i 和 U_{vi} 之间, 则需要给顾客一定的赔偿, 但对惩罚系数的要求则相对模糊, 笔者引入 CPT 通过比较配送路线上运输时间前景值来确定单位惩罚系数, 从而对配送路线进行微调. 而前景值可用价值函数和决策权重函数的乘积来表达. 单位惩罚系数表征如下:

$$\theta_{ei} = \theta_{li} = k \theta_0, \quad (4)$$

式中: θ_0 为商家确定的固定惩罚系数; k 为常数, 表示客户对时间窗要求的紧急程度.

2.1 价值函数

价值函数的确定与参考点的选取密切相关,

笔者假设客户对配送时间的预期, 参考点为 t_0 . 对于每个备选方案 G , 当 $t(G) > t_0$ 时, 任务晚于客户预期的时间到达, 表现为亏损; 当 $t(G) \leq t_0$ 时, 任务早于或吻和客户预期的时间到达, 表现为盈利或收支平衡. 则价值函数可以表示为:

$$V[t(G)] = \begin{cases} [t_0 - t(G)]^\alpha, & t(G) \leq t_0; \\ -\lambda [t(G) - t_0]^\beta, & t(G) > t_0, \end{cases} \quad (5)$$

式中, $0 < \alpha, \beta \leq 1$ 分别表征价值函数的凹凸程度, α, β 越大表示客户对收益和损失越敏感, 对运输时间越在意; $\lambda \geq 1$, 为损失相对收益的敏感性系数, λ 越大表示客户对损失的厌恶程度越高, 即对时间的要求越苛刻.

2.2 决策权重

大量研究证明: 客户在充满不确定因素的情形下进行决策时, 往往得出偏离客观概率 p 的主观概率 $\omega(p)$ 的相应结论. 根据累积前景理论可知, 主观概率大的事件其客观概率往往偏小, 即主观概率事件在某种程度上存在很大偏差.

因此 $t(G)$ 可以表示为以概率 ($p_{-m}, \dots, p_i, \dots, p_n$) 取值 ($t_{-m}, \dots, t_i, \dots, t_n$), 并且取值从大到小排列, 当 $-m \leq i \leq 0$ 时, $t_i \geq t_0$; 当 $0 \leq i \leq n$ 时, $t_i \leq t_0$.

根据 Tversky 等^[8] 在 1992 年提出的主观概率函数的形式, 决策权重计算式如下:

$$\begin{cases} \pi_n^+(G) = \omega^+(p_n), \pi_n^-(G) = \omega^-(p_{-m}); \\ \pi_i^+(G) = \omega^+(p_i + \dots + p_n) - \omega^+(p_{i+1} + \dots + p_n), \\ \quad 0 \leq i \leq n-1; \\ \pi_i^-(G) = \omega^-(p_{-m} + \dots + p_i) - \omega^-(p_{-m} + \dots + p_{i-1}), \\ \quad 1-m \leq i \leq 0. \end{cases} \quad (6)$$

$$\omega^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p^\gamma))^{1/\gamma}}; \quad (7)$$

$$\omega^-(p) = \frac{p^\delta}{(p^\delta + (1-p^\delta))^{1/\delta}}; \quad (8)$$

式中: $\omega^+(p), \omega^-(p)$ 分别为针对收益和损失时的权重函数; γ, δ 为模型参数, Kahneman 经过试验标定 $\gamma = 0.61, \delta = 0.69$.

2.3 累积前景值

在价值函数和决策权重的基础上, 根据 CPT, 路径 G 的前景值可以表示为:

$$V(G) = \sum_{i=0}^n \pi_i^+(G) V(t_i) + \sum_{i=-m}^0 \pi_i^-(G) V(t_i). \quad (9)$$

3 数学模型

在不确定情况下, 考虑不同客户对时间窗的严格要求程度的生鲜电商配送优化模型如下:

$$\begin{aligned}
\{ \min z = & \sum_i \sum_j \sum_l c_{ijl} x_{ijl} + \sum_i Q_i + \\
& \sum_i \sum_j \sum_l \alpha p x_{ijl} t_{jl}. \\
\text{s. t. } & \text{C1 } \sum_i w_i y_{li} \leq b_l \forall l; \\
& \text{C2 } \sum_l y_{li} = 1, \quad i = 1, 2, \dots, n; \\
& \text{C3 } \sum_i x_{ijl} = y_{lj}, \quad j = 0, 1, 2, \dots, n \forall l; \\
& \text{C4 } \sum_j x_{ijl} = y_{li}, \quad i = 0, 1, 2, \dots, n \forall l; \\
& \text{C5 } X_{ijl} \in s, \quad s \subseteq n; \\
& \text{C6 } s_{li}, \quad i = 0, 1, 2, \dots, n \forall l; \\
& \text{C7 } t_{ej} - (s_{ei} + t_i + t_{ij}) \leq t_{\max}; \\
& \text{C8 } t_{li} + t_{ij} \leq t_{ej}, \quad i, j = 0, 1, 2, \dots, n \forall l; \\
& \text{C9 } L_{vi} \leq s_{ei} \leq U_{vi}, \quad i = 0, 1, 2, \dots, n \forall l; \\
& \text{C10 } x_{ijl}, y_{li} = \{0, 1\}, \quad i, j = 0, 1, 2, \dots, n \forall l.
\end{aligned} \tag{10}$$

在系统优化设计问题式(10)中,目标函数表示总成本最小;约束条件 C1,使车辆 1 装载的货物总量不大于车辆的限定容量;约束条件 C2 保证每个代理点只由一辆车配送且所有代理点都得到配送;约束条件 C3、C4、C5 保证可形成回路;约束条件 C6、C7 分别是服务结束时间和最大等待时间的量化说明;约束条件 C8 表示一条线路上两邻接任务存在的条件;约束条件 C9 为时间窗约束;约束条件 C10 为规定是 0 或 1 规划问题.其中惩罚函数的惩罚因子由式(4)和(9)确定.

经多方研究证明,VRPSTW 问题属于 NP 难题.笔者所构建的模型是在 VRSTW 问题的基础上进行改进的,也归结于 NP 问题.对于这类问题人们尝试运用各种启发式算法来求解,并取得了很多成果.

由于 PSO 算法容易出现早熟收敛或停滞现象,易陷入局部极值^[9].对 PSO 算法最具代表性的几种变形有:线性递减 PSO 算法(LDPSO)^[10]、带压缩因子的 PSO 算法(CPSO)^[11].笔者结合相关文献,对基本粒子群算法在学习因子和惯性权重两方面进行改进,提高算法收敛速度.

4 粒子群优化算法与设计

4.1 PSO 算法

PSO 算法是 Kennedy 和 Eberhart 通过对鸟群飞行行为的研究于 1995 年首次提出的^[12],具有个体数目少、计算简单等优点.笔者将 PSO 应用

于生鲜电商车辆配送路径的优化求解中,取得了预期的效果. PSO 算法按式(11)、(12)来更新位置和速度,直到迭代结束.

$$\begin{aligned}
v_{ij}(t+1) = & \omega v_{ij}(t) + c_1 r_1 [p_{ij} - X_{ij}(t)] + \\
& c_2 r_2 [p_{ij} - X_{ij}(t)]; \tag{11}
\end{aligned}$$

$$X_{ij}(t+1) = X_{ij}(t) + v_{ij}(t+1), j = 1, 2, \dots, d, \tag{12}$$

式中: c_1 、 c_2 为学习因子,取值在 0 ~ 4 之间,通常 $c_1 = c_2 = 2$; r_1 、 r_2 为(0,1)之间均匀分布的随机数; ω 为惯性权重,SHI 等^[13]经过大量实验证明,如果加速度系数保持常数不变, ω 随算法迭代的进行而线性减小,将显著改善算法的收敛性能.

4.2 改进 PSO

两个学习因子在优化过程中随时间进行在 $[c_{\min}, c_{\max}]$ 进行变化,称为同步变化的学习因子,这使得粒子在优化前期能加强全局搜索能力;在优化后期有利于收敛到全局最优解.第 t 次迭代时学习因子的取值公式为

$$c_1 = c_2 = c_{\max} - \frac{c_{\max} - c_{\min}}{t_{\max}} \times t. \tag{13}$$

为克服 ω 的线性递减所带来的不足,将 PSO 算法中 ω 设定为服从某种随机分布的随机数.首先,在进化初期随机 ω 可能产生相对有效的值,加快算法的收敛速度;其次,若算法在初期找不到最优点时, ω 的随机生成可以克服 ω 线性递减使得算法最终收敛不到此最优点的局限.

ω 的计算公式为

$$\begin{cases} \omega = \mu + \sigma \cdot N(0, 1); \\ \mu = \mu_{\min} + (\mu_{\max} - \mu_{\min}) \cdot \text{rand}(0, 1), \end{cases} \tag{14}$$

式中: $N(0, 1)$ 表示标准正态分布的随机数; $\text{rand}(0, 1)$ 表示从 0 到 1 之间的随机数.

4.3 改进算法的实现步骤

同步学习因子随机权重粒子群算法的伪代码如下所示.

①Input: 种群最大迭代次数 M , 粒子数目 N 、 D 、 c_{\max} 、 c_{\min} 、 μ_{\max} 、 μ_{\min} , 以及随机权重方差

②Initialize: 初始化种群 E_0 , 设置种群代数 $g = 0$

③根据式(10)中的目标函数计算初始种群 E_0 个体适应度函数 $\text{fitness}(x(i, :))$

④For $i = 1$ to N do

⑤For $j = 1$ to D do

⑥初始化各微粒的位置 $x(i, j)$ 和速度 $v(i, j)$

⑦End for

⑧End for

⑨For $i = 1$ to N do
 ⑩ $p(i) = \text{fitness}(x(i,:))$, $y(i,:) = x(i,:)$
 ⑪End for
 ⑫ $pg = x(N,:)$
 ⑬ pg 为全局最优;
 ⑭For $i = 1$ to $(N - 1)$ do
 ⑮If $\text{fitness}(x(i,:)) < \text{fitness}(pg)$, $pg = x(i,:)$
 ⑯End if
 ⑰End for
 ⑱For $t = 1$ to M do
 ⑲根据⑬更新学习因子, ⑭更新权重
 ⑳For $i = 1$ to N do
 ㉑根据式(11)式(12)更新粒子的速度和位移
 ㉒ If $\text{fitness}(x(i,:)) < p(i)$, $p(i) = \text{fitness}(x(i,:))$, $y(i,:) = x(i,:)$
 ㉓End if
 ㉔If $p(i) < \text{fitness}(pg)$
 ㉕ $pg = y(i,:)$
 ㉖End if
 ㉗End for
 ㉘End for
 ㉙Return: 系统最小的 x_m 和 f_v

5 案例分析

由于生鲜电商的发展处于起步阶段,关于生鲜电商配送路径优化的研究及案例较少,因而采用文献[14]的案例,在 Matlab 2012a 运行环境下,对模型进行验证。该问题有 10 个配送网点,已知车速 30 km/h,每辆车的载重量为 10 t,平均运输成本为 1 元/(t·km)。群体规模为 50,学习因子最大值取 2.1,最小值取 0.8;随机惯性权重平均值的最大值取 0.8,最小值取 0.5,方差取 0.2。 $t_{\max} = 40 \text{ min}$, $\theta_0 = 20$, $Q_{\max} = 1\,000$, $\alpha = \beta = 0.88$, $\lambda = 2.25$;迭代次数 200 次,针对本问题,分别用本文算法、改进的遗传算法和传统算法进行对比,其最优解的对比结果如表 1~3 所示。

表 1 3 种算法 30 次独立运行结果比较

Tab.1 The comparison of 30 independent runs of three of algorithms

30 次运行结果	传统算法	对比算法	改进算法
最优值	4 526	4 100	3 383
最差值	4 536	4 120	3 484
平均值	4 528	4 112	3 385
方差	1.6	1.4	1.5

表 2 3 种算法成本比较结果

Tab.2 The cost comparison of three kinds of algorithm

算法	运输成本/元	货损成本/元	惩罚成本/元	违反时间窗	平均惩罚时间/min
改进算法	441	1 782	1 160	5	11.6
对比算法	430	1 800	1 880	5	12.6
传统算法	532	1 953	2 041	5	19.6

表 3 3 种算法计算结果比较

Tab.3 The calculation results of three kinds of algorithm

项目	传统算法	对比算法	改进群算法
最优解	4 526	4 100	3 383
最优解迭代次数	24	17	20
最优路线	0-2-3-9-5-1-0; 0-4-8-7-6-10-0	0-2-1-3-9-5-0; 0-4-8-7-6-10-0	0-2-1-7-9-6-0; 0-4-8-4-5-10-0

从表 2 和表 3 可以看出,笔者所采用的模型和算法在收敛速度上强于传统算法的 24 次,稍弱于对比算法的 17 次,具有如下优势:

(1) 充分考虑了顾客对时间窗的需求,尽管 3 种模型违反时间窗的个数都是 5 个,但传统算法在 5-1 的配送过程中迟到了 20 min;对比算法尽管避免了类似问题,但是在 8-7 的配送过程中,早到了 36 min,这部分惩罚成本过大;修改后平均惩罚时间和最大违反时间都有很大程度的缩减,明显缩短了顾客的等待时间。

(2) 结合时间因素在电商配送中所占的比重,在考虑惩罚成本和平均违反时间窗的情况下,本模型相比传统算法和对比算法,在运输成本和货损成本没有太明显增加的情况下,总成本节省了 1 143 元和 717 元,说明本模型更符合生鲜电商配送的要求。

(3) 根据累积前景理论来突出客户对时间窗的要求,灵活调节惩罚系数,对配送路线进行精细化处理,提升客户满意度。应用累积前景理论来检验最优路线,根据前景值的大小来判别实际送达时间与客户期望时间之间的差距,来确定单位惩罚系数是否需要调整,使配送时间更贴近客户需求,从而达到提升客户满意度的目的。

6 结论

为提高生鲜电商的配送效率,节省生鲜电商整体成本,考虑出行的不确定性因素和人们决策的有限理性特点,结合生鲜电商配送的时效性,引入前景值理论,提出了带软时间窗的生鲜电商路径优化模型。并对粒子群算法进行改进后求解该模型。根

据 CPT 理论,电商企业可以根据顾客对时间窗的要求,灵活调节惩罚因子,从而对配送线路进行微调,提升顾客对配送服务的满意度. 结合具体案例进行验证,通过对 3 种算法进行对比分析,证明了模型的有效性,求得最优路线,降低了整体成本,说明该模型在优化生鲜电商配送路径上的可行性.

参考文献:

- [1] Chinese electronic commerce research center. 2015 annual China e-commerce market data monitoring report [EB/OL]. [2017-02-16] <http://www.100ec.cn/zt/2015snda/>.
- [2] 潘璠,吴一帆,董明. 生鲜食品配送车辆路径优化[J]. 贵州农业科学,2013,41(4):223-227.
- [3] 龚树生,梁怀兰. 生鲜食品的冷链物流网络研究[J]. 中国流通经济,2006(2):7-9.
- [4] 王红玲,郑纲,何剑峰. 基于改进粒子群算法的生鲜农产品配送路径优化研究[J]. 贵州农业科学,2010,38(31):17961-17962.
- [5] SIMON H A. A behavioral model of rational choice [J]. Quarterly journal of economics, 1955, 69(1):99-112.
- [6] TVERSKY A, KAHNEMAN D. Advances in prospect theory: Cumulative representation of uncertainty [J]. Journal of risk and uncertainty, 1992, 5(4):297-323.
- [7] 任亮,黄敏,王兴伟. 考虑客户拖期厌恶行为的 4PL 路径优化问题[J]. 计算机集成制造系统,2016,22(4):1148-1154.
- [8] 张维泽,林建波,吴洪森,等. 基于改进蚁群算法的物流配送路径优化[J]. 浙江大学学报(工学版),2008,42(4):575-577.
- [9] LING S H, IU H, LEUNG F H F, et al. Improved hybrid particle swarm optimized wavelet neural network for modeling the development of fluid dispensing for electronic packaging [J]. IEEE transactions on industrial electronics, 2008, 55(9):3447-3460.
- [10] SHI Y, EBERGERT R. Empirical study of particle swarm optimized [C]//Int Conf on Evolutionary Computation. Washington: IEEE, 1999: 1945-1950.
- [11] CLERC M, KENNEDY J. The particle swarm explosion, stability, and convergence in a multidimensional complex space [J]. IEEE Transaction on evolutionary computation, 2002, 6(1):58-73.
- [12] SUN J, FANG W, Wu X J, et al. Quantum-behaved particle swarm optimization; Analysis of individual particle behavior and parameter selection [J]. Evolutionary computation, 2012, 20(3):349-393.
- [13] Y Shi, R C EBERHART. A modified swarm optimizer [C]//IEEE International Conference of Evolutionary Computation. Anchorage, Alaska: IEEE Press, 1998.
- [14] 庄景明,彭昕昀. 基于改进遗传算法的新鲜农产品配送路线优化研究[J]. 江西师范大学学报(自然科学版),2012,36(4):400-402.

Fresh Agricultural Product E-commerce Distribution Routing Problem Considering Time Demand of Customer

CUI Yan, ZHANG Zixiang, SHI Xin, WANG Xiaoliang, WANG Zhenfeng

(School of Electro-mechanical Engineering, Henan Agricultural University, Zhengzhou 450002, China)

Abstract: This paper aims to explore fresh agricultural product E-commerce routing problem under uncertain environment. Considering the characteristic of the bounded rational of customer decision, fresh agricultural product E-commerce routing optimization model was established based on Cumulative Prospect Theory (CPT). According to the characteristics of fresh agricultural E-commerce supplier for fresh features. The model took the minimum cargo damage costs, transportation costs and the penalty costs as objective, the agents point demand, goods loading and times window of logistic services about customer requirements as constraints. An improved Particle Swarm Optimization algorithm was used to solve there parts of model respectively. reference examples. Simulation and test results showed that, the model based on CPT met customer requirements on time window more accurately and improved customer satisfaction. This study could provide theoretical support to the fresh E-commerce suppliers to optimize the distribution path.

Key words: CPT; vehicle routing problem; fresh E-commerce; time windows; an improved particle swarm optimization

文章编号: 1671-6833(2017)06-0064-06

基于 Mean Shift 聚类的多级阈值化方法

程全¹, 刘晓青¹, 刘玉春¹, 王志良²

(1. 周口师范学院 机械与电气工程学院, 河南 周口 466001; 2. 北京科技大学 计算机与通信工程学院, 北京 100083)

摘要: 为了解决多级阈值化技术中所选阈值的数量通常不能预先确定的问题, 提出一种基于 Mean Shift 聚类技术的新型多级阈值化方法. 首先, 通过使用 Mean Shift 技术探寻出潜在的模式中心, 应用迭代的阈值选择方法来自动确定相邻模式中心的各个阈值; 然后, 采用多级阈值化对图像进行分割; 最后, 通过实验验证了基于 Mean Shift 聚类技术分割的图像相对于原始图像的对比度有了很大提高. 该方法通过简单修改程序参数就能够灵活控制分割精度, 可以广泛应用于单阈值分割、多级阈值分割和有损压缩等技术中.

关键词: 多级阈值化; 图像分割; 迭代阈值化; 分割质量评估

中图分类号: TP391

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.009

0 引言

图像分割是进行图像理解和目标识别的一个必不可少的步骤, 图像分割技术分为软分割和硬分割^[1]. 聚类算法是软分割方法中最主要的图像分割算法, 主要有两种: 一种是 K-means 聚类^[2], 其优点是易于实现但需要预先指定待分割图像中需要形成的群集的数量及中心; 另一种是模糊聚类^[3], 当图像中不同物体之间没有明显的边界时, 可以基于相似性准则采用模糊聚类的方法进行处理, 但是它对噪声特别敏感, 并且模糊关系的确定也非常繁琐. 软分割计算起来相当费时; 而硬分割是一种基于亮度信息的分割方法, 它假设图像的直方图含有一个或多个峰值^[4], 不需要图像的先验信息、计算简单, 实时性高, 但它忽略了图像的空间信息, 阈值选择不当, 经常会导致过分割或欠分割. 实际上, 该方法可以被认为是一个多级阈值化问题, 其中所选取的多个阈值对应于相邻显著峰值的峰谷点.

多级阈值化技术将图像的直方图分为几个部分, 当某一图像像素的特征值位于由某一对相邻阈值确定的范围时, 该像素就被分到某一相应的部分. 为了解决多阈值分割难题, 近年来群智能技术被用来寻找最优阈值. Raja 等人^[5]采用一个改进的粒子群 (PSO) 算法用于恶性肿瘤热图像的多

阈值分割. Sarkar 等^[6]基于差分进化算法 (DE) 和最小交叉熵提出了一种新颖的多阈值彩色图像分割方法.

尽管上述方法可以确定出最优阈值, 但是不能提供对应于图像直方图中潜在模式数量最优阈值数. 因此, 笔者提出了一种新型的最优多级阈值化方法, 特别适用于具有多模态直方图的图像. 首先使用 Mean Shift 程序确定出直方图的各个模式中心; 然后使用一种迭代的阈值选取方法来确定相邻模式中心的各个阈值; 最后利用已经确定的多个阈值对图像进行多阈值化分割.

1 算法部分

1.1 Mean Shift

Mean Shift 是一种非参数化的聚类方法^[7], 其最大优点是它无须预先给出聚类的数量, 且对聚类的形状没有任何限制. 作为一个迭代的模式探寻方法, Mean Shift 程序利用核来计算观察窗口的特征加权平均, 并可以准确地定位出聚类的中心并完成特征空间的划分.

对于 d 维空间 R^d 上的一个 n 点的数据集合 $\{x_i\}_{i=1, \dots, n}$, 点 x 处的多变量核密度估计为

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right), \quad (1)$$

式中: $k(x)$ 是侧面轮廓函数; h 是窗口半径. 令式

收稿日期: 2017-05-05; 修订日期: 2017-08-10

基金项目: 国家自然科学基金资助项目 (61401526); 河南省自然科学基金资助项目 (152300410134)

作者简介: 程全 (1978—), 男, 河南沈丘人, 周口师范学院副教授, 主要从事智能控制研究, E-mail: quan8888@126.com.

(1) 的梯度为0, 得到密度函数的驻点为

$$x = \frac{\sum_{i=1}^n x_i \cdot g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}, \quad (2)$$

式中: $g(x) = -k'(x)$ 是核函数 $G(x)$ 的侧面轮廓函数.

连续迭代计算 Mean Shift 矢量, 并进行移位直到程序收敛, 就能定位出局部模态. 迭代的方程为

$$y_{j+1} = \frac{\sum_{i=1}^n x_i \cdot g\left(\left\|\frac{y_j - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_j - x_i}{h}\right\|^2\right)}, j = 1, 2, \dots, (3)$$

式中: y_1 为核窗口的初始位置中心; y_{j+1} 为 y_j 处利用核 G 和窗口半径 h 计算出的加权平均.

对于 Epanechnikov 核为

$$K_E(x) = \begin{cases} (2c_d)^{-1} (d+2) (1 - x^T x), & \text{if } x^T x < 1; \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

式中: c_d 为单位 d 维球体的体积.

Mean Shift 的公式为

$$y_{j+1} = \frac{1}{n_{y_j}} \sum_{y_j x_i \in S_h(y_j)} x_i - y_j, \quad (5)$$

式中: $S_h(y_j)$ 为中心位于 y_j ; 半径为 h ; 包含 n_{y_j} 个数据点的超球体.

选择图像像素的灰度值作为特征空间, Mean Shift 程序可以探寻出其潜在的概率密度模态. 算法的步骤描述如下:

(1) 将特征空间划分为 n 个非重合的同等大小部分. 每一部分的大小为 $(G_{\max} - G_{\min})/n$, 其中 G_{\max} 和 G_{\min} 分别为图像像素灰度值的最大值和最小值. 为了避开低密度区域, 要保证每一部分中像素的数量不少于一个阈值 T_1 .

(2) 运行 Mean Shift 程序 n 次以得到 n 个收敛点, Mean Shift 程序和核半径为 $h = (G_{\max} - G_{\min})/n$.

(3) 将距离小于一个预设阈值 T_2 的相邻收敛点合并为一点, 求出 m ($m < n$) 个潜在的概率密度模态中心.

为了将图像灰度分为 m 个类, 比如 $0 \sim 255$, 考虑到图像灰度范围的上下限已确定, 需要再确定 K 个阈值, 满足 $K = m - 1$. 在下文中采用一种迭代阈值选取方法来确定 K 个阈值.

1.2 多级阈值化

为了使分割方法更具鲁棒性, 阈值应该能

够自动地被确定出来. 因此, 自动阈值选择是图像分割的一个关键步骤. Ridler and Calvard^[8]提出了一种迭代的阈值选择方法, 该方法描述如下.

给定一幅图像, 假定 $I(i, j)$ 是 (i, j) 处像素的灰度值, 其中 $1 \leq i \leq M$ and $1 \leq j \leq N$, $0 \leq I(i, j) \leq L - 1$, L 为图像的灰度级别. 首先计算出图像的灰度直方图, 假定希望能找到一个阈值 T 并利用该阈值将图像分割为一个二进制图像. 在二进制图像中, 所有灰度值低于 T 的像素, 其灰度值都被 A 所取代; 所有灰度值高于 T 的像素, 其灰度值都被 B 所取代. 将误差函数定义为二进制图像和原始图像相应像素灰度值差的平方和. 使用积分以方便表达, 则误差函数可以表示为:

$$e^2 = \int_0^T (k - A)^2 h(k) dk + \int_T^{L-1} (k - B)^2 h(k) dk, \quad (6)$$

式中, k 为像素的灰度值; $h(k)$ 为灰度值在直方图中出现的次数. 分别令 e^2 对 T 、 A 和 B 的偏导数为 0 可以得到:

$$T = \frac{A + B}{2}; \quad (7)$$

$$A = \frac{\int_0^T k \cdot h(k) dk}{\int_0^T h(k) dk} = \mu_1; \quad (8)$$

$$B = \frac{\int_{T+1}^{L-1} k \cdot h(k) dk}{\int_{T+1}^{L-1} h(k) dk} = \mu_2. \quad (9)$$

实际上, A 和 B 分别为直方图被阈值 T 分成的两部分的均值.

注意到阈值 T 仅由两部分的均值 A 和 B 确定, 而这两部分的均值 A 和 B 又只有当 T 确定了以后才能被计算出来. 所以需要使用迭代算法: 首先选定初始阈值做为起始点; 然后利用此阈值将直方图分为两个部分, 分别计算出两部分的均值, 并将阈值更新为两部分均值的和的一半, 该过程反复执行直到阈值收敛; 最后利用多阈值化方法图像的灰度范围将被 K 个阈值分割为 $K + 1$ 个部分.

给定原始图像 $I(i, j)$, 利用 Mean Shift 程序确定阈值的数量 K , 令 $T_{\min} = 0$ 及 $T_{\max} = L - 1$, 利用迭代阈值选取方法计算出 K 个阈值 $\{T(i) | i = 1, 2, \dots, K\}$, 其中 $T(i) < T(i + 1)$. K 个阈值将图像直方图分割为 m 个互不重叠的区域:

$$J(i, j) = \begin{cases} 0, & T_{\min} \leq I(i, j) \leq T(1); \\ \frac{L-1}{K+1}, & T(1) < I(i, j) \leq T(2); \\ \frac{2 \cdot (L-1)}{K+1}, & T(2) < I(i, j) \leq T(3); \\ \dots & \\ L-1, & T(K) < I(i, j) \leq T_{\max}, \end{cases} \quad (10)$$

其中, J 为分割图像。

2 分割质量评估

图像分割质量的评估没有一个标准的方法,最常用的是主观评估,这种方法主观性太强,而非监督性评估方法由于可以克服上述缺点,受到越来越多的关注. Haralick 等^[9]给出了高质量图像分割结果的四条准则:①分割出的各个区域内部像素的特性要尽量一致;②相邻区域之间像素特性的差异要大;③区域内部没有孔洞;④区域边界在空间位置上要精确,并且不能支离破碎。

前两条被用于审查图像中目标的特性,被称为特性准则;后两条被称为语义准则,被用来衡量人如何将一个区域识别为一个目标^[10]。

Zhang^[11]引入期望区域熵作为区域内一致性的度量.假定原始图像被分割成 N 个区域,采用 R_j 来表示区域 j ($1 \leq j \leq N$) 中的像素集合; S_j 表示区域 j 的面积, S_{jm} 表示对应于区域 j 的原始图像中具有灰度值 m ($0 \leq m \leq L-1$) 的像素个数,那么区域 j 的熵定义为

$$H(R_j) = - \sum_{m=0}^{L-1} \frac{S_{jm}}{S_j} \log \frac{S_{jm}}{S_j}. \quad (11)$$

而分割的期望区域熵为

$$Hr(I) = \sum_{j=1}^N \frac{S_j}{S_I} H(R_j), \quad (12)$$

式中: S_I 是图像 I 的面积,表示图像内部具有更高的相似度。

Biswas 等^[12]提出了用 IQI (image quality index) 来度量区域间的差异. 对于一个行数和列数分别为 M 和 N 的图像 I , 令 $I(i, j)$ 表示 (i, j) 处像素的灰度值. 则 IQI 被定义为

$$IQI = \frac{\sum_{i=1}^M \sum_{j=1}^N c(i, j)}{M \times N - \sum_{i=1}^M \sum_{j=1}^N h(i, j)}; \quad (13)$$

$$c(i, j) = \frac{|B - I(i, j)|}{B}; \quad (14)$$

$$h(i, j) = \frac{1}{8} \sum_{\substack{p=i-1, \\ p \neq i}}^{i+1} \sum_{\substack{q=j-1, \\ q \neq j}}^{j+1} \exp(-|I(i, j) - I(p, q)|). \quad (15)$$

IQI 反映的是沿着区域间边界图像像素性质的平均对比度,值越大表明图像的平均对比度越大。

3 实验结果

选择两组复杂度不同的图像来验证该多阈值图像分割方法的有效性. 第一组图像是一幅合成图像“Gray8”;第二组图像包含从 Berkeley Segmentation Dataset^[13]选出来的自然图像 35 010、126 007、241 004、113 016、385 039、368 016. 笔者提出的多阈值分割方法包括 3 个参数:核半径 h 控制着分割的敏感度;阈值 T_1 暗示了特征空间中最小可以接受的概率密度;阈值 T_2 对应于相邻两个收敛点的最小距离. 实验选取 $h = T_2 = (G_{\max} - G_{\min})/10$, 其中, G_{\max} 和 G_{\min} 分别为图像像素灰度值的最大值和最小值; $T_1 = 0.001 \times (M \times N)$, 其中 M 和 N 分别为图像的行列数。

图 1 和图 2 给出了两组图像的分割结果. 中间图像显示了直方图分割的过程示意图,矩形表示整个灰度直方图中被划分成的几个互不重叠的部分,每个矩形的中心是其 Mean Shift 程序核窗口的起始点,每一个五角星代表了相应 Mean Shift 程序的收敛点,如果相邻收敛点的距离小于核窗口的半径则将此相邻收敛点合并以得到模态中心,各模态中心用黑色三角形表示,相邻模态中间的阈值由迭代阈值选取方法选定并由短划线表示其位置。

图 1 所示的 8 个离散的灰度值被成功地检测出来(图 1(b)),并给出了比较精确的分割结果(图 1(c)). 对于图 2 中的 2 幅自然图像,也给出了令人满意的分割结果,并且相对于原始图像,分割图像的对比度提高了许多。

采用本文方法得到的图像的分割结果如图 3 所示,从左到右分别为“Lena”“Cameraman”和“Baboon”;其中第一行显示了原始图像,第二行为笔者所提多阈值分割方法得到的图像分割结果,其中合理选择分割半径后得到 $K=6$ 个分割阈值。

将本文分割方法与典型的多阈值分割方法 PSO 及 DE 算法进行了比较. 分割图像的性能参数选择前文所述的期望区域熵 Hr 和图像质量指标 IQI ,比较结果见表 1. 可以看出,对于对比度比较明显的自然图像,选择同样数量的阈值数是时,

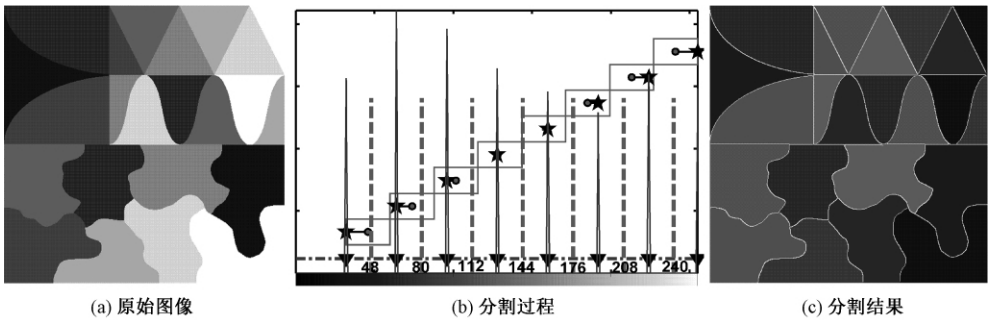


图 1 Gray8 图像分割情况
Fig.1 Gray8 image segmentation

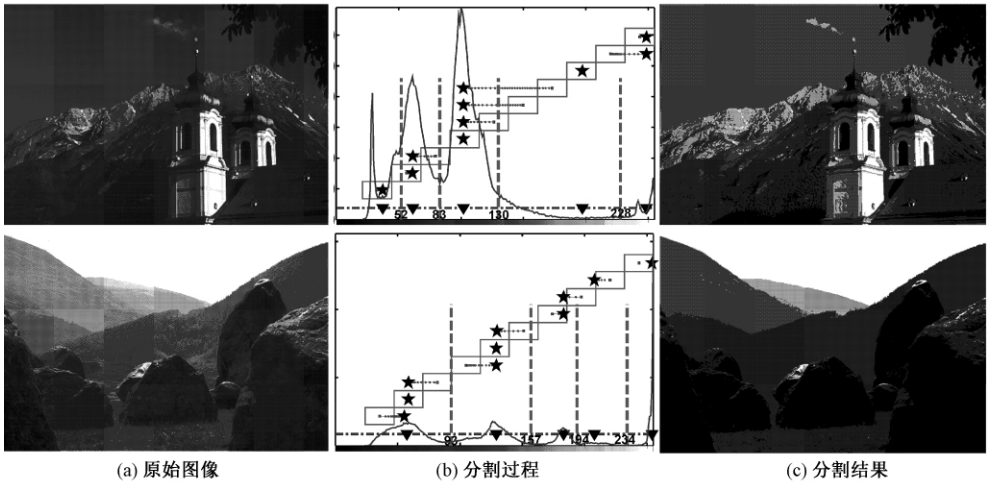


图 2 自然图像分割过程及结果

Fig.2 Segmentative process and result of image segmentation

表 1 3 种方法的期望区域熵和图像质量指标比较

Tab.1 Expected region entropy and image quality index for three method

<i>Hr</i>			<i>IQI</i>		
本文方法	PSO	DE	本文方法	PSO	DE
6.258	6.176	5.984	1.159	1.073	1.177
9.730	9.663	9.274	1.843	1.470	1.732
4.911	4.997	4.773	0.438	0.457	0.461

本文方法的性能优于 PSO 和 DE 算法.但是对于纹理较丰富的图像如 Baboon,本文方法稍逊于 PSO 和 DE 算法.

本文分割算法在 Berkeley Segmentation Dataset 测试数据上获得了良好的分割效果.为了展示本文分割方法的性能,如图 3 所示,选取其中的几幅典型图像 35 010、126 007、241 004、113 016、385 039、368 016,实验选取分割阈值个数为 $K=6$.

与各群智能算法一样,本文方法也具有随机性.表 1 和表 2 为几种方法的最优结果.分割算法对每幅图像重复运行 50 次后,分别计算目标函数的均值和方差,结果如表 2,3 所示,可以看出本文的图像分割算法具有良好的稳定性.

表 2 3 种方法的图像分割期望区域熵性能稳定性比较

Tab.2 Expected region entropy performance stability of image segmentation for three method

测试图像	本文方法		PSO		DE	
	均值	标准差	均值	标准差	均值	标准差
Lena	6.303	0.618	6.190	1.874	6.017	0.976
Cameraman	9.679	0.590	9.472	0.489	9.305	0.732
Baboon	5.017	1.803	5.135	1.309	4.931	2.390
35010	11.370	0.210	11.408	0.430	11.140	0.790
126007	5.943	0.255	5.810	2.320	5.798	2.018
241004	6.374	0.387	6.193	1.358	6.158	2.041
113016	5.179	0.068	5.104	0.289	5.092	1.034
385039	7.385	0.189	7.272	1.487	7.490	1.376
368016	6.941	0.576	6.577	1.965	6.782	1.069

对于 $K=6$,将本文方法的计算时间与 PSO 和 DE 方法进行了比较,结果如表 4 所示.结果显示,大部分情况下本文算法运行时间最少.

笔者采用的多阈值化方法用到了 3 个参数:核半径 h 、阈值 T_1 和阈值 T_2 . T_1 和 T_2 对分割结果的影响要远小于 h .一般来说,若选取较大的 T_1 值,则 Mean Shift 程序只能探测出最重要的模



图3 原始测试图像分割的比较

Fig.3 The compare of original test images and segmented images

表3 3种方法的图像分割 IQI 性能稳定性比较
 Tab.3 Image quality index performance stability of image segmentation for three method

测试图像	本文方法		PSO		DE	
	均值	标准差	均值	标准差	均值	标准差
Lena	1.148	0.765	1.079	1.494	1.157	1.376
Cameraman	1.870	0.613	1.526	0.421	1.772	1.232
Baboon	0.419	1.490	0.441	1.689	0.450	1.660
35010	0.519	0.471	0.528	0.474	0.535	0.609
126007	1.280	0.598	1.197	2.501	1.206	2.103
241004	1.102	0.608	1.097	1.169	1.137	1.942
113016	0.739	0.236	0.728	0.376	0.699	1.284
385039	1.363	0.160	1.309	1.790	1.475	1.018
368016	2.031	0.376	1.947	1.483	1.962	1.215

表4 3种方法的计算时间比较
 Tab.4 Computation time for three method s

测试图像	本文方法	PSO	DE
Lena	1.485	1.871	2.138
Cameraman	1.995	2.395	2.017
Baboon	1.938	1.850	1.749
35010	2.116	2.287	2.431
126007	1.649	1.927	1.818
241004	2.017	2.251	2.514
113016	1.729	1.954	2.137
385039	1.652	1.798	1.819

态,这通常会造成欠分割;而若选取较小的 T_1 值,则 Mean Shift 程序还能探测出次要的模式,这通常会造成过分割.实际上, T_1 选取为整个图像大小的 1% 就可以得到较好的自然图像分割结果.

而 T_2 对分割结果影响较小,通常令 T_2 等于核窗口的半径,因此我们主要讨论 h 的影响.

作为一种非参数化的密度估计方法, Mean Shift 程序的估计精度会受到核半径 h 的影响;通常较小的核半径值会使程序得到太多模式;而较大的核半径值会生成平滑的密度和较少数量的模式.最常用的方法是选择不同的核半径,将算法运行多次直到分割性能达到最优. Comaniciu^[14] 用比较复杂的方式解决了核半径的选择问题,实验显示,当核半径选为 $(G_{\max} - G_{\min}) / 10$ 时,就可以得到较好的分割结果,分割图像的对比度提高了许多.可以看到,随着核半径的增加, H_i 越来越小,表明随着图像被划分的区域数目的增多,各个区域内像素特性的差异越来越小; IQI 也越来越小,表明区域间像素特性的差异越来越小,尽管如此,任何一个分割结果的 IQI 都要比原始图像的 IQI 大.因此,从图像压缩的角度来看,本文方法也具有一定的参考价值.

4 结论

为了解决多级阈值化图像分割方法中阈值的数量不能预先确定这一难题,提出了一种新型的基于聚类技术的多级阈值化方法.首先使用 Mean Shift 技术搜索到潜在的模式中心点;然后应用迭代的阈值选择方法来确定相邻模式中心的各个阈值;最后采用多级阈值化方法对图像进行分割,并通过大量实验得到了验证.此外,还应用期望区域熵 H_r 和图像质量指标 IQI 对分割结果进行了评估,显示了本文方法的有效性.本文方法可以通过简单地修改程序参数来控制分割的精度,并且无

须预先确定直方图中潜在的模态数量,因此,本文方法可应用于单阈值分割、多级阈值分割和有损压缩等技术中。

参考文献:

- [1] TOKAS N, KARKRA S, PANDEY M K. Comparison of digital image segmentation techniques a research review [J]. International journal of computer science and mobile computing, 2016(5) : 215 – 220.
- [2] DHANACHANDRA N, MANGLAM K, CHANU Y J. Image segmentation using K-means clustering algorithm and subtractive clustering algorithm [J]. Procedia computer science, 2015, 54: 764 – 77.
- [3] Gómez D, Yáñez J, GAUDA C. Fuzzy image segmentation based upon hierarchical clustering [J]. Knowledge based system, 2015, 87(c) : 26 – 37.
- [4] OTSU N. A threshold selection method from grey-level histograms [J]. IEEE transaction on systems man & cyberntics, 2014, 9(1) : 62 – 66.
- [5] RAJA N S M, SUKANYA S A, NIKITA Y. Improved PSO based multi-level thresholding for cancer infected breast thermal images using otsu [J]. Procedia computer science, 2015, 48: 524 – 529.
- [6] SARKAR S, DAS S, CHAUDHURI S S. Amultilevel color image thresholding scheme based on minimum cross entropy and differential evolution [J]. Pattern recognition letters, 2015, 54: 27 – 35.
- [7] CHENG Y. Mean shift, mode seeking, and clustering [J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 17(7) : 790 – 799.
- [8] RIDLER T W, CALAARD S. Picture thresholding using an iterative selection method [J]. IEEE transactions on systems man & cybernetics, 2010, 8(8) : 630 – 632.
- [9] HARALICK R, SHAPIRO L. Survey: image segmentation techniques [J]. Computer vision, graphics and image processing, 2013(29) : 100 – 132.
- [10] 余金煌, 陶同赞. 小子域滤波在高密度电法图像处理中的应用 [J]. 水电技术, 2015(1) : 5 – 10.
- [11] ZHANG H, FRITTS J E, GOLDMAN S A. A region entropy-based objective evaluation method for image segmentation [C] // 2009 IEEE Instrumentation and Measurement Technology conference. Singapore: IEEE 2014: 38 – 49.
- [12] BISWAS S, LOVELL B C. Bezier and splines in image processing and machine vision [M]. London: Springer, 2009.
- [13] Berkeley segmentation dataset [EB/OL]. [2017-03-01]. [http://www. cs. berkeley. edu/projects/vision/bds/](http://www.cs.berkeley.edu/projects/vision/bds/). 2012.
- [14] COMANICIU D. An algorithm for data-driven bandwidth selection [J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 25(2) : 281 – 288.

Based on the Mean Shift Clustering Multilevel Threshold Method

CHENG Quan¹, LIU Xiaoqing¹, LIU Yuchun¹, WANG Zhiliang²

(1. Mechanical and Electrical Engineering, Zhoukou Normal University, Zhoukou 466001, China; 2. School of Computer & Communication Engineering, Beijing University of Science & Technology, Beijing 100083, China)

Abstract: In order to solve the problem that the number of selected thresholds in multilevel thresholds cannot be usually predetermined, a novel multi-level thresholding method based on Mean Shift Clustering technique was proposed. Using Mean Shift technology to explore the potential mode center, the various thresholds of adjacent to the mode center was automatically determined by using of iterative threshold selection method, and then the method of multi-level threshold was used for image segmentation. The experimental results showed that, relative to the original image, contrast of the image split with Mean Shift clustering technique was greatly improved. This method could control the segmentation precision flexibly by simply modifying parameters of the program, and could be widely used in the technology of single threshold segmentation, multi-level threshold segmentation and detrimental compression and other technologies.

Key words: multilevel thresholding; image segmentation; iterative threshold selection; segmentation evaluation

文章编号: 1671-6833(2017)06-0070-04

石油钻井工程预警技术发展概述

李 广^{1,2}, 张 衡¹, 王 杰¹, 朱晓东¹, 岳彩通¹

(1. 郑州大学 电气工程学院, 河南 郑州 450001; 2. 中国电子科技集团公司 第二十二研究所, 河南 新乡 453001)

摘 要: 钻井工程预警技术是钻井安全的关键保障, 通过井场实时钻井过程参数监测、海量钻井数据挖掘与智能学习、异常状态预警模型建模与优化、特征量异常在线判断等过程, 实现石油钻井异常状态预警, 预防钻井工程事故发生. 通过介绍钻井工程预警技术架构并比较其特点, 从而对钻井工程预警技术未来发展进行展望.

关键词: 石油钻井; 事故预警; 数据挖掘; 特征量提取; 建模

中图分类号: TE277

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.010

0 引言

石油钻井是一项复杂的地下工程^[1-2]. 钻井过程伴随大量随机、模糊、不确定和不可控因素. 由地层结构及水文、钻井工艺及设备、油气分布及钻井环境等各种原因造成的钻井事故, 严重威胁着钻井安全. 因此, 实时监测钻井状态, 准确预报钻井事故, 对于保障钻井安全、降低钻井成本和提高钻井效益具有重要意义. 笔者旨在回顾钻井工程预警技术的研究历程, 通过讨论当前钻井工程预警技术的整体架构和所采用的技术, 比较各类预警系统的特点及不足, 总结钻井工程预警技术的现状及存在的问题, 从而探索石油钻井事故预警系统相关理论和技术的研究新方向和新方法.

1 石油钻井预警技术发展简介

上世纪七十年代之前, 钻井工程预警主要靠人工进行, 主观判断钻井过程中是否有事故发生或即将发生, 需要很强责任心和很丰富的经验. 直到本世纪初, 石油钻井工程预警才升级为设置软件参数门限值进行预报, 但参数门限设置的主观性依然很大, 依然受限于现场专家的责任心和经验.

近十多年, 随着传感器技术^[3]、数据挖掘技术^[4]、人工智能技术^[5]和压缩感知技术(compressive sensing, CS)^[6]的发展, 预警系统智能程度越来越高. 国内外公司投入大量资金开发相关技术和软件, 如国外哈里伯顿公司 RTOC 实时监测技术; 国内中国电子科技集团公司第二十二研究所开发的钻井工程预警系统. 虽然国内外公司在钻井预警技术方面投入大量资金, 但目前钻井预警技术应用效果仍不理想. 国外依据邻井资料建立预测模型, 当实钻井与邻井的地层分布差别很大时, 需要实时修正预警模型, 导致误报和漏报较多. 国内预警系统模型采用固定和定制方式, 模型固定适用性差, 误报率和漏报率高, 定制模型需要专家经验, 人为因素影响大.

2 钻井工程预警技术

钻井工程预警技术包含传感器采集技术、数据预处理技术、滤波技术、数据挖掘技术、特征量提取技术和多参数融合技术等.

2.1 信号采集及预处理

信号预处理包括信号失真判断和滤波处理, 一般方法有 3σ 方法, 又称拉伊达方法. 它先假定

收稿日期: 2017-04-02; 修订日期: 2017-07-10

基金项目: 国家自然科学基金资助项目(61473266)

作者简介: 李广(1980—), 男, 河南商丘人, 郑州大学博士研究生, 主要从事智能计算、机器学习等方面研究.

通信作者: 王杰(1959—), 男, 河南周口人, 郑州大学教授, 博士生导师, 主要从事定量反馈理论、人工智能及数据挖掘等研究, E-mail: wj@zzu.edu.cn.

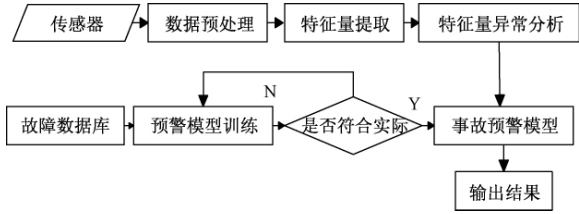


图 1 预警系统框图

Fig. 1 Diagram of pre-warning system

一组只含有随机误差的数据,对其计算得到标准偏差,按一定概率确定一个区间.认为区间内的数据是正常数据,区间外的数据是粗大误差,含有该误差的数据应予以剔除或着重处理.

图 2(a) 是立管压力参数带噪声数据,采用 3σ 方法滤波后的图形如图 2(b) 所示,野点剔除是系统正确预警的前提.

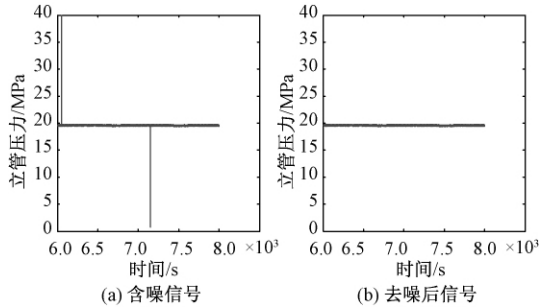


图 2 带噪声数据和剔除野点后数据

Fig. 2 Noise and after eliminating offending points

2.2 特征量提取

特征量即由原始值映射到其他空间的值,如变化率、振幅值和短期均值等.对实时系统数据,一般采用加矩形窗的方式对时间序列数据进行截断.

(1) 均值计算如图 3 所示,矩形框是滑动窗,所有的计算都在矩形窗内进行.均值线通过动态计算滑动窗内数据得到,上下阈值线内的数据认为是正常数据,上阈值与上边界值和下阈值与下边界值认为是粗大误差,需要剔除或重点处理.



图 3 均值计算算法示意图

Fig. 3 Schematic diagram of the mean value calculation

(2) 振幅值计算方法:先求均值;然后计算当前值与均值误差绝对值幅值.

(3) 变化率计算如图 4 所示.为避免变化率的计算受到个别采样点的影响,该时间段两端的参数数据取其邻域的平均值,也就是上图中所示

的均值时间段.

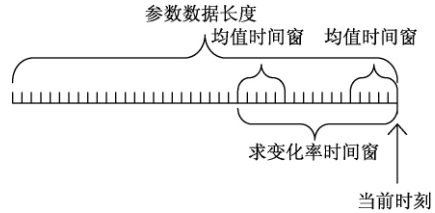


图 4 变化率算法计算示意图

Fig. 4 Schematic diagram of velocity of variation calculation algorithm

2.3 数据融合技术

数据融合^[7]一般分为像素级融合、特征级融合和决策级融合.像素级融合即原始参数的融合;特征级融合是指从各原始数据中提取特征,然后融合这些特征;决策级融合的对象是决策结果,原始数据源都经过特征量提取,获得一定准则和决策的可信度信息,最后对参数的属性决策结果进行融合,最终得到唯一决策.

表 1 溢流井漏模型参数关联表

Tab. 1 Parameters table of overflow leakage model

项目	总池体积	出口流量	入口流量
溢流	上升	上升	无异常
井漏	下降	下降	无异常

2.3.1 D_S 证据推理技术融合

D_S 证据理论^[8]是采用处理互补信息和不确定信息的数据融合算法.文献[5]把神经网络初步诊断的结果看成 D_S 融合的证据,避免了 D_S 证据理论中概率分配函数的构造,使得融合后的信息含量比融合前大得多,降低了事故诊断的不确定性.文献[2]把神经网络输出经归一化之后作为 D_S 算法的证据,该方法集成多层次状态监测与异常状态诊断模型,不仅实现了当前状态多传感器信息融合,还实现了当前状态与历史状态的证据融合.

由于 D_S 方法存在以下因素使得应用效果并没有想象中理想:①计算上存在潜在的组合爆炸问题;②无法处理高度冲突的证据.当证据完全冲突时,无法使用 D_S 理论进行融合;③要求证据必须是独立的,在某些情况下并不满足这样的条件;④证据合成规则没有非常坚实的理论支持.

2.3.2 神经网络技术融合

人工神经网络 (artificial neural networks, ANNs)^[9-11].文献[9]针对卡钻采用 BP 神经网络构建预测模型,针对 BP 算法收敛速度慢、学习效率低等缺陷,采用自适应学习速率法对神经网络进行改进,提高了神经网络的预测精度,减少了网

络训练所需时间. 文献[10]建立了基于 Monte-Carlo 模拟神经网络和的钻井工程风险评估方法. 文献[11]利用神经网络超强并行处理问题能力和学习能力, 建立多个针对井漏在不同钻井状态下的诊断模型, 最后利用证据论合成规则进行融合. 人工神经网络由于梯度算法收敛慢, 同时存在着局部最小问题, 不能够保证钻井事故诊断的实时性和准确性. 并且在故障诊断中最重要的是对知识的获取和经验的利用, 而对于钻井作业的特殊性, 神经网络的训练样本却难以获取, 而且如何从有限的故障样本中得到具有较大推广能力的决策函数又是另外一道难题, 以至于神经网络在钻井事故诊断方面的应用受到限制.

2.3.3 模糊理论技术融合

模糊推理^[12-14]是由 Zadeh 首先提出, 是模拟人日常推理的一种近似推理. 笔者前期建立了分层模糊推理的预警系统, 解决了模糊推理中, 参数过多造成模糊规则爆炸的问题. 文献[12]采用模糊推理理论, 构建了钻井工程事故预警系统, 并通过现场实际数据的仿真实验表明: 推理结果与现场实际相符, 并通过构建多种事故模型知识库, 使预警系统有较大的灵活性和适应性. 文献[13]提出了一种针对溢流, 采用基于 AHP 思想的模糊专家系统的预警模型. 能够解决针对钻井过程中无法获知邻井资料和地质信息的前提下, 无法建立基于样本信息的溢流预警模型的问题. 模糊推理虽然模拟了人类的一般推理过程, 但是由于模糊推理过程中参数过多, 往往会造成维数灾的问题, 严重影响系统的实时性. 虽然文献[14]提出了解决此问题的方法, 但是增一型分层模糊推理在多参数推理过程中的误差大小还没有理论方面的研究与支持.

2.3.4 灰色关联技术融合

灰色关联理论^[15-16]是一种研究不确定信息和小样本的方法. 文献[15]采用灰色聚类法建立安全评价模型, 将现场采集的有关钻井安全数据应用于灰色聚类钻井安全评价模型中, 对油气钻井作业存在的安全情况进行评价和验证. 文献[16]通过多目标智能灰靶理论、层次分析以及成分分析相结合的思路, 给出了相应的计算方法, 建立钻井工程事故综合预警模型, 提出五种事故类型预警模型和等级划分结果.

由于钻井工程预警模型复杂多样, 采用单一 GM 模型不能很好地适用所有的模型, 此外, 背景值对于指数型增长序列具有较好的模拟预测效

果, 对于线性变化序列、有异常数据序列以及非等距序列等问题仍需要进一步的研究.

2.3.5 支持向量机技术融合

支持向量机(SVM)^[17-18]在实际应用中, 存在的一个突出问题是: 如何选取算法中的关键参数, 因为不同参数组合决定了其学习性能和推广能力的大小. 目前大多数的解决思路是融合其他的算法来选择参数. 文献[17]提出了 SVM 和 PSO 相结合的钻井事故预测模型, 并通过实例应用验证了该预警模型的有效性. 文献[18]提出了基于 RS-SVM(粗糙集-支持向量机)的人工智能诊断方法, 先利用粗糙集理论对样本数据进行属性约简, 不改变目标分类的基础上消去冗余特征, 达到降低样本维数的目的, 最后将处理后的数据输入到支持向量机进行训练.

SVM 本质上是分类算法, 效果取决于分类目标边界可分, 但对于钻井工程中模糊多变的目标函数适用性还需要进一步研究.

2.3.6 时间序列技术融合

时间序列^[19-20](AR、MA、AR-MA、ARIMA 等模型). 文献[19]针对钻井过程中卡钻事故, 提出了基于时间序列的卡钻预测方法, 通过构建的时序模型对钻井数据进行预处理, 对各个 ARMA 模型做功率谱估计作为预测卡钻的重要参考. 文献[20]提出了 ARIMA 模型, 以风险概率作为时间序列的原始值, 提取统计特征, 定量预测钻井作业现场未来发生风险的概率, 经比对, 该模型短期预测比长期预测效果好.

虽然时间序列在工程中广泛应用, 并取得了良好的效果, 但是对瞬间发生的工程事故复杂, 时间序列分析效果不理想. 很多工程事故的发生, 与参数变化的时间顺序有较强的关联关系, 但目前的文章中均没有提及此类问题的分析.

3 总结与展望

随着超深井、水平井、大位移井等出现, 钻井过程中出现风险的可能性在增加, 因此亟需一套功能强大的预警系统为钻井工程保驾护航. 以下是可能的发展方向:

1) 传感器技术发展. 传感器井下化是发展趋势, 技术难题包括传感器在恶劣井下环境条件下(如高温高压)的承受能力、传感器供电问题、传感器安装位置是否影响正常钻进等方面.

2) 井下数据传输技术. 目前正在兴起的压缩感知技术把压缩放在了硬件, 直接采样压缩后的

数据,在地面接收到数据后,通过重构算法重构井下原始数据。井下数据传输技术是未来发展方向。

3) 数据挖掘技术的发展。油田保存了大量事故数据,而这些数据均没有得到很好的开发和利用。从海量故障数据里,寻找事故相关联的知识,增强预警系统的智力水平,是未来发展的一个研究方向。

4) 多参数融合技术。随着智能技术发展,在提高模型预测精准度,模型的适应度,算法的鲁棒性等方面,是多参数融合技术研究的一个方向。

参考文献:

- [1] 王杰,李广,朱晓东. 基于分层模糊推理的石油钻井事故预警系统[J]. 微计算机信息, 2008, (21): 177-178+184.
- [2] 廖明燕. 钻井过程实时状态监测与诊断技术研究[D]. 青岛:中国石油大学机电工程学院, 2010.
- [3] YU Y, HUANG S Y, WANG J, et al. Design of wireless logging instrument system for monitoring oil drilling platform[J]. IEEE sensors journal, 2015, 15(6): 3453-3458.
- [4] WU X D, ZHU X Q, QING G, et al. Data mining with big data[J]. IEEE transactions on knowledge and data engineering, 2014, 26(1): 041-4347.
- [5] 樊帆. 基于多源信息的钻井事故预警系统研究[D]. 西安:西安石油大学石油工程学院, 2013.
- [6] CANDLES E, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE transactions on information theory, 2006, 52(2): 489-509.
- [7] KHALEGHI B, KHAMIS A, KARRAY F O, et al. Multisensor data fusion: a review of the state-of-the-art[J]. Information fusion, 2013, 14(1): 28-44.
- [8] 姜延吉. 多传感器数据融合关键技术研究[D]. 哈尔滨:哈尔滨工程大学计算机科学与技术学院, 2010.
- [9] 朱丹. 基于神经网络的卡钻预测方法研究[D]. 西安:西安石油大学电子工程学院, 2013.
- [10] 魏凯,管志川,韦杰宏,等. 基于神经网络和 Monte-Carlo 模拟的钻井工程风险评估方法[J]. 中国安全科学学报, 2013, (2): 30-33.
- [11] 蔡汶君. 基于神经网络融合技术的钻井井漏诊断模型研究[D]. 成都:西南石油大学计算机科学学院, 2014.
- [12] 吴俊杰. 钻井工程事故监测和预警方法研究[J]. 录井工程, 2006, 17(1): 53-55.
- [13] 张禾,李祁颖,张露之,等. 基于模糊专家系统的钻井溢流智能预警技术[J]. 西南石油大学学报, 2016, 38(2): 169-175.
- [14] 朱晓东,王杰. 基于分层模糊系统的石油钻井参数预测模型[J]. 石油报, 2010, 31(5): 838-848.
- [15] 谭露. 基于灰色聚类的钻井作业安全评价方法的研究与实现[D]. 成都:西南石油大学计算机科学学院, 2014.
- [16] 王建彬. 石油钻井工程事故的预警技术研究[D]. 北京:中国地质大学安全工程系, 2013.
- [17] 孙万海. 基于 SVM 和 PSO 的钻井事故智能预警系统研究[D]. 西安:西安石油大学计算机学院, 2014.
- [18] 连晓圆. 钻井过程中故障检测与诊断方法的研究[D]. 大连:大连理工大学电气工程学院, 2013.
- [19] 陶宇龙. 基于时间序列的卡钻预测方法研究[D]. 西安:西安石油大学电子工程学院, 2014.
- [20] 赵春兰,王兵,郭平. 基于 ARIMA 模型的钻井作业风险预测[J]. 新疆石油地质, 2013, 34(4): 469-472.

A Review: Pre-warning System of Oil-Drilling Engineering

LI Guang^{1,2}, ZHANG Heng¹, WANG Jie¹, ZHU Xiaodong¹, YUE Caitong¹

(1. School of Electrical Engineering Zhengzhou University, Zhengzhou 450001, China; 2. 22th Research Institute of China Electronics Technology Group Seventh Research Department, Xinxiang 453001, China)

Abstract: Warning technology of drilling engineering was the key technology of drilling safety protection. Through the monitoring of real-time well site drilling process parameters, huge amounts of drilling data mining and intelligent learning, abnormal state modeling and optimization, abnormal characteristics of the early warning model online judging process, achieved the goal of oil drilling abnormal state early warning, and prevention of drilling engineering accidents. This paper reviewed the development course of early warning technology, introduced the drilling engineering warning technology architecture, and also introduced the early warning technology in detail and compared their characteristics, finally depicted the development of future early warning system for drilling engineering.

Key words: drilling; pre warning; data mining; feature extraction; modeling

具备直流故障隔离能力的新型 MMC 拓扑研究

张清枝¹, 李志辉²

(1. 新乡学院 机电工程学院, 河南 新乡 453003; 2. 郑州大学 电气工程学院, 河南 郑州 450001)

摘要: 对柔性直流输电的直流侧短路故障问题, 提出了一种新型的能够自我阻断直流故障的子模块拓扑, 该拓扑可以通过自身结构特性来阻断故障电流; 在此基础上, 为了抑制故障时子模块电容电压升高带来的不利影响, 提出了具有阻尼电阻的子模块拓扑, 该拓扑在阻断故障电流的同时还具有抑制子模块电容电压的作用. 在 PSCAD/EMTDC 平台下搭建了 9 电平 MMC-HVDC 仿真模型, 并验证了所提出新型拓扑的正确性.

关键词: 柔性直流输电; 直流故障; MMC; 子模块拓扑; 子模块电容电压

中图分类号: TM46; TM721.1 **文献标志码:** A **doi:** 10.13705/j.issn.1671-6833.2017.06.011

0 引言

近年来, 随着分布式电源、储能技术的不断发展, 直流输电技术受到了人们的广泛关注^[1-3]. 柔性直流技术可以独立控制功率、没有换向失败, 成为目前研究的热点之一. 而模块化多电平换流器 (multilevel modular converter MMC) 由于采用模块化设计, 具有可拓展性好, 开关频率低等优点^[4-5], 目前已成为换流站的首选拓扑. 在采用 MMC 的系统中, 直流侧的短路故障是目前 MMC 需要研究的一个主要课题^[6-7]. 实际工程中通常采用 MMC 半桥型子模块 (half bridge sub module, HBSM), 当换流器直流侧发生双极短路故障时, 在 IGBT (insulated gate bipolar transistor) 关断的情况下, 短路电流会经过并联的反向二极管流入故障点, 且无法通过闭锁换流器来阻断短路电流, 会造成十分严重的危害^[8].

采用可以自我阻断直流故障电流的 MMC 拓扑是目前较为可行的方法. 该方法通过换流器本身的结构特点, 通过故障时二极管两端的反向电压来阻断直流故障电流. 由于 IGBT 的开关速度很快, 因此能够确保切断故障的迅速性. 并且系统能够快速恢复非永久性故障, 提高了供电可靠性.

笔者利用 SM 换流器拓扑的结构特征提出了一种改进型拓扑, 可以在不改变控制/调制方法的前提下, 使系统具备处理直流故障的能力, 并且对故障时的子模块充电电压加以抑制, 减小故障带来的危害^[9-10]. 根据直流故障的条件, 在 PSCAD/EMTDC 环境下进行了时域仿真.

1 MMC 基本结构及子模块拓扑

1.1 MMC 的基本结构

MMC-HVDC 系统主要由 SM 电路装置级联来实现^[11-12]. 目前广泛采用的是基于半桥型 SM 的 MMC. 三相半桥型 MMC 的结构单元拓扑如图 1 所示.

图 1 中, 左侧 A、B、C 三相接入交流电网, 右侧输出直流电压. MMC 每一相拥有上和两个桥臂, 其桥臂电压和电流分别为 U_{arm_PA} 、 I_{arm_PA} 、 U_{arm_NA} 、 I_{arm_NA} . 每一个完整的上桥臂或下桥臂均由 N 个 SM 上下级联再与一个桥臂外端的电抗器 L_0 串联组成, U_{dc} 表示直流侧的输出电压, O 表示电压为 0 的参考点. 图 1 的底部为一个 SM 的半桥型拓扑结构, 其输出的子模块电容电压为 U_{SM} ; 通过子模块的电流为 i_{arm} ; 电容器两端的电压为 U_C ; T_1 、 T_2 为控制子模块运行状态的 IGBT; D_1 、 D_2 为与之并联的反向二极管.

收稿日期: 2017-04-07; 修订日期: 2017-07-11

基金项目: 国家自然科学基金资助项目 (61473266); 2016 年度河南省高等学校重点科研项目 (16A470013)

作者简介: 张清枝 (1970—), 女, 河南新乡人, 新乡学院副教授, 主要从事电力电子技术及自动化运行控制方面的研究, E-mail: 13937395104@163.com.

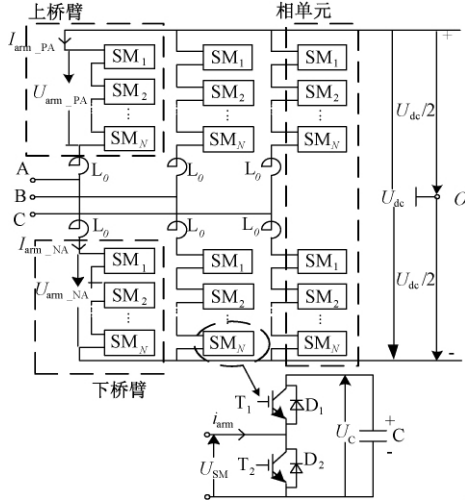


图1 MMC的基本结构

Fig.1 MMC basic structure

1.2 直流故障情况下的子模块拓扑分析

在直流侧短路故障期间,半桥型 SM 电路,如图 2 所示. 故障发生时,所有的 IGBT 均被阻断,直流故障电流通过与 T_2 并联的反向二极管 D_2 从交流侧流出. 由此可见,半桥型 SM 不具有直流故障阻断能力.

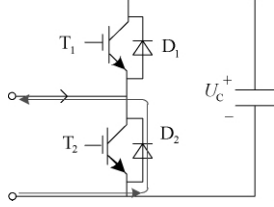


图2 半桥型 SM 拓扑结构

Fig.2 HBSM topology

全桥型 SM 电路如图 3 所示,全桥的 MMC-HVDC 系统的功率损耗以及成本比半桥型更高. 直流故障时,所有 SM 的 IGBT 被阻断,电容器可以产生反向电压阻挡交流侧的电流,从而为全桥型 SM 提供直流故障处理能力.

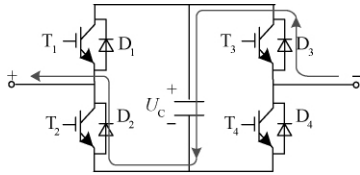


图3 全桥型 SM 拓扑结构

Fig.3 FBSM topology

2 新型拓扑的结构及故障清除原理

2.1 新型拓扑的基本结构

结合 FBSM 的直流故障处理能力,笔者对传统的 HBSM 进行改进,提出了一种新型子模块拓扑,如图 4 所示. 这种子模块拓扑在半桥型和全桥

型子模块拓扑的基础上进行融合,与全桥型 SM 相比,改进的拓扑多用了—个限流电阻 R_1 ,省去了—个 IGBT.

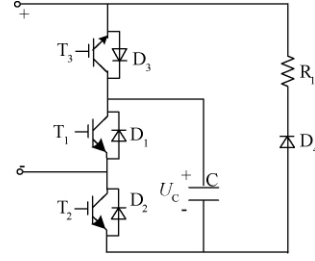


图4 改进的子模块拓扑结构

Fig.4 The improved SM topology

2.2 新型的故障清除原理

正常运行时, T_3 始终处于导通状态,子模块通过 T_1 和 T_2 的轮换投切来改变其输出的电压为 0 或 U_c .

子模块的电容电压与交流电压满足:

$$\begin{cases} U_{dc} = NU_c; \\ U_{ph} = mU_{dc}/2; \\ U_L = \sqrt{3}U_{ph}, \end{cases} \quad (1)$$

式中, U_{ph} 、 U_L 分别为交流侧的相电压和线电压幅值; m 为调制比; U_c 为子模块电容电压; N 为导通的子模块个数.

如图 5 所示,短路电流经 $D_1 \rightarrow C \rightarrow D_4 \rightarrow R_1$ 流出子模块. 同理,MMC-HVDC 发生双极短路故障时,两相之间的短路电流共流过 $2N$ 个 $D_1 \rightarrow C \rightarrow D_4 \rightarrow R_1$ 这样的组合以及 2 个电抗器 L .

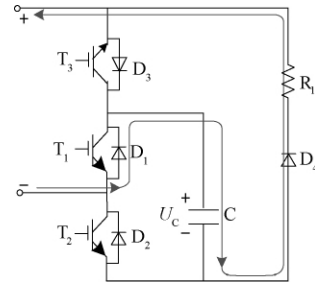


图5 故障时的短路电流通路

Fig.5 Short-circuit current paths in DC fault

以 A、B 相为例,根据 KVL(基尔霍夫电压定律):

$$u_{ba} = [N(u_{D1} + u_{D4} + U_c + u_{R1}) + u_L] \times 2 = [N(u_{D1} + u_{D4} + u_{R1}) + U_{dc} + u_L] \times 2, \quad (2)$$

式中, u_{ba} 和 u_L 分别为交流线电压和电抗器电压的瞬时值; u_{D1} 、 u_{D4} 为二极管的电压瞬时值; u_{R1} 为限流电阻 R_1 上的电压瞬时值.

整理得:

$$u_{D1} + u_{D4} = \left(\frac{u_{ba}}{2} - u_L - U_{dc} \right) / N - u_{R1}. \quad (3)$$

显然, $u_{D1} + u_{D4} < 0$, 即二极管 D_1 和 D_4 承受反向电压, 因此可以达到切断故障电流的效果。

2.3 新型拓扑的故障控制策略及作用

图 5 所示, 发生直流双极短路故障时, 二极管 D_1 、 D_4 上的压降可以忽略, 因此相当于串联电阻 R_1 与子模块电容 C 进行串联分压, 可得

$$U_c = \frac{U_{dc}}{N} \times \frac{\frac{1}{j\omega C}}{\frac{1}{j\omega C} + R_1} = \frac{U_{dc}}{N(1 + j\omega CR_1)} \quad (4)$$

可以看出, 发生直流故障时, 随着阻尼电阻的增大, 子模块电容充电电压随之减小, 故而阻尼电阻起到抑制直流电容电压的效果。

故障时控制策略如图 6 所示, 当系统发生直流故障时, 所有的 IGBT 将要闭锁, 以阻断故障电流。对于永久性故障, 需要断开交流断路器以隔离故障并进行修复; 对于瞬时性故障, 故障电流被切断之后, 需要解锁 IGBT 以重新建立直流电压, 系统恢复正常运行状态。由于电力电子器件的动作十分迅速, 且故障清除过程不需要交流断路器的动作, 因而系统能够快速恢复。

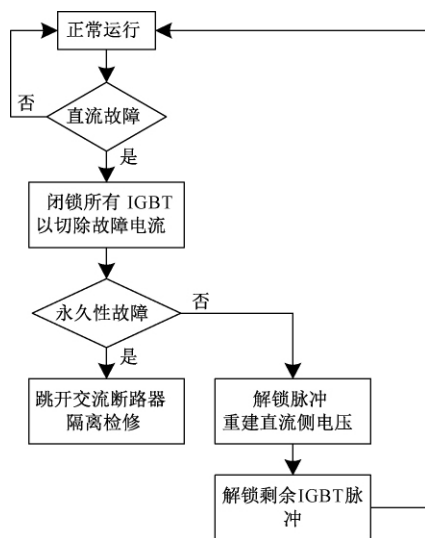


图 6 直流故障的控制策略

Fig. 6 DC-fault control strategy

采用上述控制策略, 故障清除时间能够控制在 10 ms 之内, 可以有效保护子模块中的 IGBT 和二极管。因此, 所提的新型子模块拓扑能够显著提升柔性直流输电系统的故障阻断及恢复能力。

3 仿真验证与分析

3.1 直流故障阻断能力验证

为了验证 SM 拓扑能够抑制直流故障电流, 在 PSCAD/EMTDC 软件平台下搭建了单端 9 电

平 MMC-HVDC 模型。

模型的具体参数为: 直流电压 ± 320 kV; 传输功率 1 000 MW; 系统运行频率为 50 Hz; 变压器变比为 230/352 kV, 漏电抗为 0.15 p.u.; 每个上/下桥臂由 4 个新型子模块和 4 个 HBSM 串联; 桥臂电抗为 112 mH; 子模块电容及电压为 164 μ F 和 80 kV。系统采用定功率控制, 在直流侧设置了双极短路故障, 进行仿真分析。设置故障发生在 $t = 1.10$ s, 故障持续时间 1 s。

子模块在 0.2 s 前预充电, 在 0.9 s 之后, 当发生双极短路故障时, 直流电流超过 3.12 kA, 此时所有的 IGBT 闭锁。

MMC-HVDC 的仿真运行逻辑为, 当 1.10 s 发生双极短路故障时, 直流电流迅速冲击到 3.12 kA, 使所有的 IGBT 闭锁, 在故障持续时间内, 若直流电流快速下降为 0, 则说明拓扑具有直流故障阻断能力。

如图 7 所示, 根据时序控制, 0.2 s 时子模块充电完毕, 充电电源断开, 直流电流从 0 开始上升, 0.6 s 到达 1.5 kA, 1.10 s 发生直流侧双极短路故障, 直流电流在 1.1007 s 达到 3.12 kA, 使所有的 IGBT 闭锁, 闭锁后直流电流迅速下降为 0, 说明该模型具有直流故障阻断能力。图 8 为该系统的直流电压波形, 正常运行时维持在 640 kV, 发生双极短路故障后电压下降为 0。MMC 子模块中的 IGBT 两端电压波形如图 9 所示。

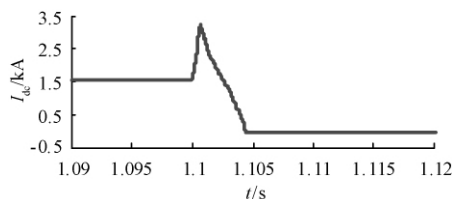


图 7 直流电流波形

Fig. 7 DC current waveform

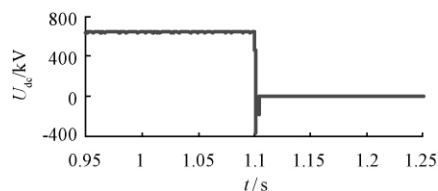


图 8 直流电压波形

Fig. 8 DC voltage waveform

3.2 子模块电容充电过程分析

首先在工频情况下, 设置 $R_1 = 0$, 得到短路故障发生时三相桥臂电流如图 10 所示。由图 10 可以看出, 1.1007 s 时, IGBT 闭锁, 经过 $T_{\sigma_1} = 0.8$ ms, A、B 两相的上桥臂电流衰减至零。后经过 $T_{\sigma_2} =$

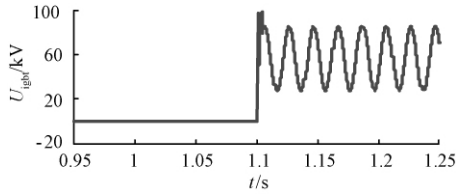


图9 IGBT 两端的电压波形

Fig. 9 IGBT voltage waveform

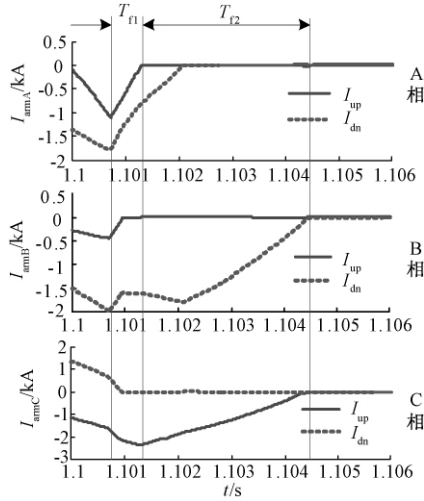


图10 三相桥臂电流波形(0 Ω)

Fig. 10 Three-phase arm current waveform(0 Ω)

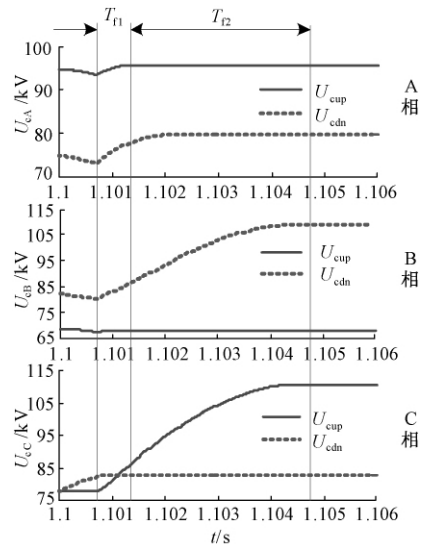


图11 三相子模块平均电容电压波形(0 Ω)

Fig. 11 Three-phase SM average capacitance voltage waveform(0 Ω)

3.2 ms,三相所有的桥臂电流均降为零,故障电流完全被阻断.在 T_{f1} 所对应的故障时间内,A相的上下桥臂、B相下桥臂及C相上桥臂子模块开始充电,如图11所示.当A、B相不再有短路电流通过后,A相上桥臂子模块电容充电完成.在三相短路电流均被阻断后,所有的子模块电容充电完成.可以看出,B相下桥臂和C相上桥臂的子模块电

容电压从80 kV 上升到近115 kV,上升率过高,从而会导致系统的不稳定.

3.3 阻尼电阻对充电电压的抑制仿真验证

为解决上述问题,采用具有阻尼电阻的改进新型拓扑来抑制子模块电容电压的升高.由于直流线路电感的储能,在故障时会使 U_{dc} 产生反向衰减电压,考虑故障时直流电压的反向冲击不大于额定值的2倍,因此设定 $R_1 = 150 \Omega$,得到三相桥臂电流和子模块电容电压如图12~13所示.

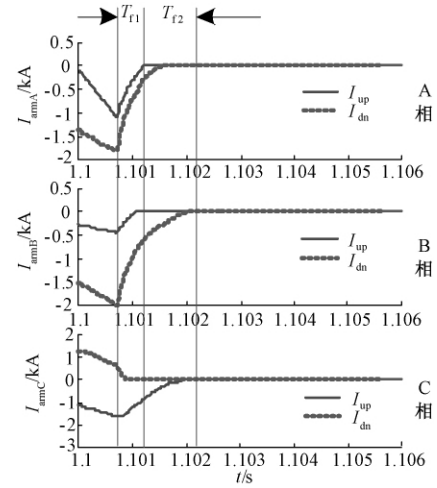


图12 三相桥臂电流波形(150 Ω)

Fig. 12 Three-phase arm current waveform(150 Ω)

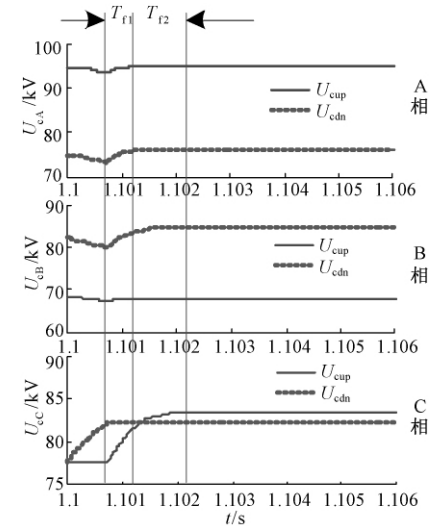


图13 三相子模块平均电容电压波形(150 Ω)

Fig. 13 Three-phase SM average capacitance voltage waveform(150 Ω)

对比图10、12可以看出, $R_1 = 150 \Omega$ 的改进拓扑在故障发生后,其短路电路的阻断时间由4 ms减小为1.5 ms,反应更为迅速;对比图11、13可以看出,各相子模块的电容充电电压上升的趋势得到了有效减缓,尤其B、C两相的充电电压由原先的110 kV降低为85 kV左右,得到了很好的抑制.

4 结论

针对传统 HBSM 不能解决直流故障的问题,笔者结合半桥型和全桥型 SM 的特点,提出了一种新型的子模块拓扑.该拓扑具备阻断直流故障电流能力的同时还能够抑制子模块电容过电压.相比于其他能够阻断直流故障短路电流的拓扑,新型改进拓扑结构简单,应用电力电子器件数量少,还能够有效减小短路瞬间造成的危害.

最后在单端 9 电平 MMC-HVDC 混合仿真模型中,对不同 R 值的直流侧双极短路故障下的子模块电流及电压进行对比.仿真结果表明所提出的 SM 拓扑能迅速切断故障电流,并且能够很好地抑制故障时子模块电容的充电电压.

参考文献:

- [1] QIN J, SAEEDIFARD M, ROCKHILL A, et al. Hybrid design of modular multilevel converters for HVDC systems based on various submodule circuits [J]. IEEE transactions on power delivery, 2015, 30(1): 385–394.
- [2] 刘军会. 轻型高压直流输电仿真研究与实验装置设计[D]. 郑州: 郑州大学电气工程学院, 2014.
- [3] 张灵娟. 灵宝换流站交直流侧故障仿真分析[D]. 郑州: 郑州大学电气工程学院, 2012.
- [4] 蔡新红, 赵成勇, 庞辉, 等. 向无源网络供电的 MMC-HVDC 系统控制与保护策略[J]. 中国电机工程学报, 2014, 34(3): 405–414.
- [5] 仇雪娜, 赵成勇, 庞辉, 等. 基于 MMC 的多端直流输电系统直流侧故障控制保护策略[J]. 电力系统自动化, 2013, 37(15): 140–145.
- [6] 张建坡, 赵成勇. MMC-HVDC 直流侧故障特性仿真分析[J]. 电力自动化设备, 2014, 34(7): 32–37.
- [7] ADAM G P, AHMED K H, FINNEY S J, et al. New breed of network fault-tolerant voltage source converter HVDC transmission system [J]. IEEE transactions power system, 2013, 28(1): 335–346.
- [8] 李笑倩, 刘文华, 宋强, 等. 一种具备直流清除能力的 MMC 换流器改进拓扑[J]. 中国电机工程学报, 2014, 34(36): 6389–6397.
- [9] 汪波, 胡安, 唐勇, 等. IGBT 电压击穿特性分析[J]. 电工技术学报, 2011, 26(8): 145–150.
- [10] 刘建平, 侯凯, 李伟邦, 等. 串联 IGBT 保护模块设计和故障诊断分析[J]. 电气传动, 2015, 45(10): 76–80.
- [11] 向往, 林卫星, 文劲宇, 等. 一种能够阻断直流故障电流的新型子模块拓扑及混合型模块化多电平换流器[J]. 中国电机工程学报, 2014, 34(29): 5171–5179.
- [12] 丁云芝, 苏建徽, 周建. 基于钳位双子模块的 MMC 故障清除和重启能力分析[J]. 电力系统自动化, 2014, 38(1): 97–103.

Research on the Topology of MMC with DC Fault Isolating Capability

ZHANG Qingzhi¹, LI Zhihui²

(1. College of Mechanical and Electrical Engineering, Xinxiang University, Xinxiang 453003, China; 2. School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: To solve the MMC-HVDC dc-fault, a new sub-module topology with DC faults blocking capability was proposed, which could be used to block the fault current by its structure characteristic. On this basis, in order to restrain the adverse effects of sub-module capacitor voltage rising during the fault time, a new sub-module topology with a damping resistor was proposed. The topology not only could block fault current, but also had sub-module capacitor voltage suppression effect. An 9 level SBMMC-HVDC simulation model was built in PSCAD/EMTDC platform, and the simulation proved the proposed new topology was correct.

Key words: flexible HVDC; dc faults; MMC; sub module topology; sub module capacitor voltage

路面免碾压贫混凝土基层工作性研究

盛燕萍¹, 孙仕伟¹, 赵澍林¹, 周辉丽², 陈拴发¹

(1. 长安大学 材料科学与工程学院, 陕西 西安 710064; 2. 青海省交通科学研究院, 青海 西宁 810008)

摘 要: 设计了一种免碾压贫混凝土基层, 它可有效保证贫混凝土基层的稳定性和平整度, 同时可减少施工机械、减化施工工序、缩短周期、改善施工环境. 通过研究免碾压贫混凝土基层材料, 分析了影响贫混凝土工作性的因素, 并提出采用维勃稠度试验作为其工作性评价方法, 以及维勃稠度作为其工作性评价指标, 最后, 采用多因素正交设计试验, 并根据极差分析和方差分析, 得到评价指标维勃稠度的回归公式. 结果表明, 维勃稠度值能够较好地评价免碾压贫混凝土的工作性.

关键词: 道路工程; 贫混凝土; 免碾压; 维勃稠度值; 工作性

中图分类号: U416 217

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2016.06.014

0 引言

目前, 半刚性基层已是我国高等级公路的主要结构类型, 其强弱和好坏显著影响整个路面结构的使用质量和使用寿命^[1]. 贫混凝土(Lean Concrete)是水泥用量较少的混凝土, 相比水泥稳定碎石和石灰碎石等半刚性基层, 其承载力更高、抗冲刷性能和耐久性更好, 是一种优质的基层材料^[2-3].

现阶段重载交通路面结构破坏的现场调查显示, 损坏很大程度上是由于承载力造成, 部分因为基层施工原因导致的温缩和干缩裂缝进而形成路面结构反射裂缝. 若调整基层施工方式和工艺, 提高基层的承载力, 能够改善路面的使用性能, 延长使用寿命^[4]. 目前, 具备优良性能的贫混凝土基层已经得到广泛应用, 但现场施工大部分依然采用传统的振动成型或插捣成型工艺, 针对贫混凝土水泥含量少的特殊性, 笔者提出了不同于普通碾压方式的免碾压成型方法, 成型过程中无振动和插捣步骤, 依靠材料自身的嵌挤和粘附让混合料达到密实性. 免碾压与传统振捣施工相比, 可以显著提高贫混凝土基层的质量, 减少和优化施工机械和工序, 减少施工时间, 降低施工噪音对环境的污染. 笔者对免碾压多孔混凝土进行了较为深

入的研究^[5-10], 但多孔混凝土不同于贫混凝土, 笔者结合免碾压贫混凝土的特点, 重点对其工作性进行分析, 提出了针对免碾压贫混凝土基层的工作性评价方法和指标.

1 免碾压贫混凝土工作性影响因素

工作性至少包含流动性和黏聚性两个部分, 但这两个方面又相互矛盾. 混凝土增加流动性的过程是混合料中自由水增加的过程, 必将导致混合料的黏聚性减小; 而为保证混合料的黏聚性和整体性, 不管降低用水量还是增加增稠介质, 都会限制混合料的自由水量, 降低其流动性. 因此, 混凝土在施工时要根据结构形式和现场条件, 选择流动性和黏聚性最佳的混合料组合^[11-13].

1.1 集浆比

保持贫混凝土材料的水灰比不变, 加入的水泥浆量越多, 则集料颗粒表面的包裹浆层将越厚, 润滑效果更好, 混凝土具有更好的流动性, 反之则小. 若混合料中浆量过多, 流浆和泌水现象将不可避免, 降低了贫混凝土的粘聚性和保水性, 影响其强度与耐久性. 若混合料中水泥浆量太少, 集料表面得不到充分裹覆, 会减小贫混凝土的黏聚性, 甚至产生崩塌, 因此, 贫混凝土的集浆比应以满足拌

收稿日期: 2016-07-28; 修订日期: 2017-01-19

基金项目: 国家自然科学基金资助项目(51208047; 51208046), 中国博士后科学基金资助项目(2013M532005; 2015T81000), 青海省科技计划项目(2017-22J-763; 2017-25ZJ-764)

作者简介: 盛燕萍(1981—), 女, 浙江临安人, 长安大学副教授, 主要从事道路材料与结构方面的研究, E-mail: shengyanping2003@163.com.

合物流动性要求为准。

1.2 水灰比

确定集浆比后,水灰比将决定混合料中的水泥浆稠度。水灰比小说明水泥浆稠,混合料的流动性将较小;反之,水灰比较大,虽然混合料具有较大的流动性,但其整体性和保水性将变差。因此,在实际工程中若通过增加用水量提高混合料流动性时,水泥用量必需同时加大,确保水灰比保持恒定。

1.3 砂率

集料总表面积和空隙率受到砂率变化的影响,且水泥砂浆不仅会润滑粗骨料,还将粗集料之间从点接触变成面接触,显著提高集料间的粘结效果,改善贫混凝土基层的工作性,提高其强度和稳定性。

2 工作性评价方法和指标

贫混凝土的水泥胶结料含量较少,强度主要靠粗集料彼此之间的内摩阻力形成,细集料填充空隙,对免碾压贫混凝土而言,既要保证混合料有足够的流动性,又要满足混合料之间的粘聚性而确保强度。免碾压贫混凝土工作性研究分析的目的是寻找保证良好工作性的水泥砂浆量,以此评价其工作性。

坍落度试验和维勃稠度试验是目前普遍采用的检测混凝土稠度的方法。对于免碾压贫混凝土,坍落度试验并不能准确测出不同拌合物稠度的差异性;对于高流动性混合料,坍落度数值更不能反

映其成型后稳定性的优劣,且稠度差异的混合料也依然可能具有相同的坍落度值,即坍落度大小与混合料稠度的高低并不存在对应关系。

维勃稠度试验^[14]的操作过程很短,可以评价干硬性混合料填充模型的能力,弥补坍落度试验对测试低流动性混合料灵敏度的不足,虽然维勃稠度试验对于流动性高,坍落度大的新拌混凝土灵敏度差,但由于免碾压贫混凝土相对普通成型的混合料的浆集比更高,在现场连续施工时的可操作性更强,因此,可以采用维勃稠度评价免碾压贫混凝土的工作性。

3 免碾压贫混凝土工作性评价

3.1 试验用原材料

笔者所用水泥为秦岭 42.5 强度等级的普通硅酸盐水泥,技术指标见表 1。所用粗细集料均为西安本地石灰岩碎石,技术指标分别列于表 2 和表 3。采用规范推荐的级配,具体级配要求见表 4。

维勃稠度值受多方面因素的影响,笔者采用正交试验进行试验安排。采取含水量(水占水泥、砂和集料总质量的百分比)、水泥剂量及砂率 3 个因素,各因素根据实践经验取如表 5 所示的 4 个水平。试验选择正交表 $L_{16}(4^5)$ 的前 3 列安排试验,其中 4、5 列为空列。根据该试验安排,具体的配合比和测试结果见表 6。在此基础上,对正交试验结果分别进行极差分析和方差分析。

表 1 水泥技术指标

Tab.1 Technical index of cement

项目	水泥胶砂弯拉强度/MPa		水泥胶砂抗压强度/MPa		凝结时间/min		细度/%	标准稠度需水量/%	安定性
	3 d	28 d	3 d	28 d	初凝	终凝			
实测值	3.95	8.7	19.3	45.8	232	318	3.76	28.0	合格

表 2 粗集料技术指标

Tab.2 Technical index of thick aggregate

试验项目	松堆密度/ ($\text{g} \cdot \text{cm}^{-3}$)	吸水率/%	含泥量/%	针、片状含量/%	压碎值/%
实测值	1.452	0.877	0.40	5.82	6.50

表 3 细集料技术指标

Tab.3 Technical index of thin aggregate

试验项目	细度模数	含泥量/%	表观密度/ ($\text{g} \cdot \text{cm}^{-3}$)	吸水率/%
实测值	2.86	2.10	2.937	1.283

表 4 粗集料级配范围要求

Tab.4 Thick aggregate gradations

筛孔/mm	31.5	26.5	19	16	9.5	4.75	2.36
通过筛孔百分率/%	100	95 ~ 100	60 ~ 75	30 ~ 50	10 ~ 30	0 ~ 10	0 ~ 5

表 5 贫混凝土配合比试验正交设计因素水平表

Tab.5 Table of factors and levels of orthogonal design for lean concrete mix ratio test

水平 s	因素 r		
	A	B	C
	含水量 w /%	水泥剂量 c /%	砂率 β_s /%
I	5.75	7	33
II	6.00	8	36
III	6.25	9	39
IV	6.50	10	42

3.2 极差分析

对表 6 所示试验结果进行极差分析,其结果如表 7 所示。表 7 中 I_j 是指第 j 列取 I 水平下的指标和, R_j 是指第 j 列中 I_j , II_j , III_j 和 IV_j 之间最大值与最小值之差,称为极差。极差的大小代表该

因素变化对指标值的影响程度.通过极差的大小排序可以得到各个影响因素对 VB 值的影响大小,其相应顺序如表 8 所示.

在此基础上,得到各个影响因素的不同水平值与相应指标的变化趋势关系,如图 1 所示.

表 6 正交设计和配合比安排与试验结果

Tab.6 Orthogonality design, mixture ratio arrangement and test results

配比 编号	试验号	A	B	C	D 空列	E 空列	配合比/($\text{kg} \cdot \text{m}^{-3}$)				维勃稠度 VB/s
							水	水泥	砂	集料	
I-1	1	1	1	1	1	1	113	128	604	1227	54
I-2	2	1	2	2	2	2	108	139	627	1115	51
I-3	3	1	3	3	3	3	106	152	660	1032	65
I-4	4	1	4	4	4	4	111	175	734	1014	80
II-1	5	2	1	2	3	4	112	123	630	1121	25
II-2	6	2	2	1	4	3	113	139	575	1168	31
II-3	7	2	3	4	1	2	111	153	715	987	42
II-4	8	2	4	3	2	1	112	170	664	1038	49
III-1	9	3	1	3	4	2	117	122	682	1067	22
III-2	10	3	2	4	3	1	116	138	723	999	19
III-3	11	3	3	1	2	4	119	157	576	1169	32
III-4	12	3	4	2	1	3	118	172	619	1101	29
IV-1	13	4	1	4	2	3	123	123	741	1023	29
IV-2	14	4	2	3	1	4	125	142	692	1083	20
IV-3	15	4	3	2	4	1	124	157	629	1118	29
IV-4	16	4	4	1	3	2	126	176	581	1180	17

表 7 极差计算结果

Tab.7 Results of range analysis

维勃稠度	A	B	C
I_j	250	130	134
II_j	147	121	133
III_j	102	168	156
IV_j	95	175	170
R_j	155	54	36
最优方案	A4	B2	C2

一般规律,无法得出对 VB 有显著影响的因素.因此,对其进一步采取方差分析,并通过分析结果归纳回归相应的关系式,为评价免碾压贫混凝土基层的工作性提供定量化依据,根据表 6 的试验结果进行方差分析,结果如表 9.

从表 9 可以看出,当置信率为 95% 时,含水量对贫混凝土的维勃稠度有显著影响;含水量和水泥用量均会显著影响贫混凝土的维勃稠度,该影响因素的排序和表 7 中的极差分析结论一致.

对正交表的极差分析仅能得到对 VB 影响的

表 8 各因素对指标的影响及最佳因素水平组合

Tab.8 The factors influence on the index and the best factors combination

考核指标	各因素对考核指标影响大小的排序	最佳因素水平组合
维勃稠度 VB	A(含水量) > B(水泥用量) > C(砂率)	A4、B2 和 C2

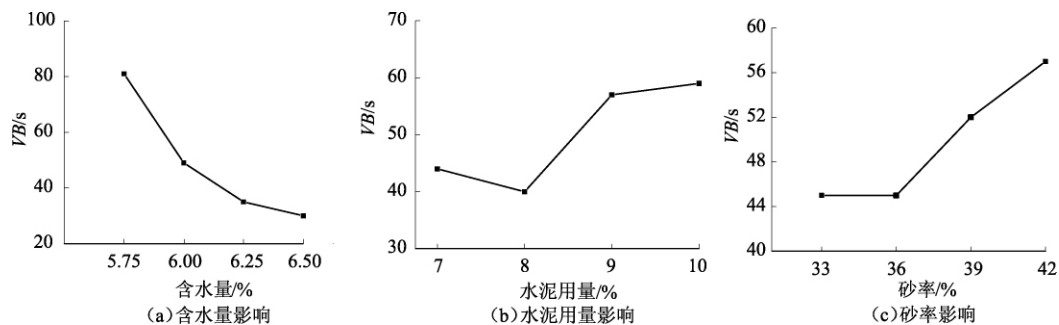


图 1 各个因素的水平变化影响

Fig.1 The influence of various factors on the level of change

表 9 正交表的方差分析

Tab.9 Variance analysis of orthogonal arrays

方差来源	离差平方和 SS	自由度 DF	方差 MS	F 值	临界值	显著性
A(含水量)	3 800.48	3	1 266.83	25.32	$F_{0.95}(3,6) = 4.76$	显著
B(水泥用量)	537.36	3	179.12	3.58		
C(砂率)	245.06	3	81.69	1.63		
D(空列)	205.12	6	50.04	—		
E(空列)	95.12	6	50.04	—		

3.3 提出回归方程

对表 6 中数据进行回归分析,得到维勃稠度 VB 与含水量、水泥剂量及砂率的关系:

$$VB = 268.514 - 5.084.3w + 452.975c + 110.492\beta_s, \quad R = 0.889. \quad (1)$$

式中: VB 为维勃稠度, s; w 为含水量, %; c 为水泥剂量, %, 水泥剂量 $c = \text{水泥质量} / (\text{砂质量} + \text{石子质量})$; β_s 为砂率, %.

另外,施工工艺也会明显影响免碾压贫混凝土的工作性,不同施工方式应满足表 10 要求的不同维勃稠度值^[13].

表 10 不同施工方式下贫混凝土的维勃稠度推荐值

Tab.10 Vebe consistometer recommended value of lean concrete in different construction ways s

摊铺方式	滑模摊铺机	轨道摊铺机	三辊轴机组
出机维勃稠度	15 ~ 35	25 ~ 45	35 ~ 55
摊铺维勃稠度	20 ~ 40	30 ~ 50	40 ~ 60

对于提出的回归方程(1),其参数和对应结果是在较为理想的室内条件下得出,而贫混凝土在实际施工过程中,受气候、温度、降水等诸多环境因素影响,会导致维勃稠度值产生不规律波动,为进一步平滑该波动,笔者在提出的维勃稠度值基础上,增加允许波动时间 5 s,以此平滑 VB 的变动和切合现场情况,即免碾压贫混凝土工作性的允许波动范围为 $VB \pm 5$ s,若实测值在此范围内,则认为贫混凝土满足工作性要求.

4 结论

(1) 贫混凝土免碾压成型方式充分利用材料的自密实性,利于提高混凝土强度和降低混合料渗透性,能够减少施工机械和施工工序及周期,对提高生产率,减少噪音和环境保护有积极的效果.

(2) 贫混凝土是干硬性混凝土,其组成材料、水灰比、集浆比和砂率是影响贫混凝土工作性的主要因素,通过正交试验设计得出各因素对工作性影响的顺序:含水量 > 水泥用量 > 砂率.

(3) 维勃稠度(VB)是能够简单、易行、快速

评价免碾压贫混凝土工作性的适宜指标,通过试验研究得出免碾压贫混凝土满足工作性要求的允许波动范围为 $VB \pm 5$ s.

参考文献:

- [1] 盛燕萍,李海滨,陈拴发.基于抗冲刷性能的半刚性基层材料组成设计方法[J].郑州大学学报(工学版),2012,33(3):19-23.
- [2] 王浩,陈剑,谭华.贫混凝土基层水泥混凝土路面结构使用性能对比[J].中外公路,2011,31(5):46-50.
- [3] 付其林,陈拴发,彭种.贫混凝土基层沥青路面温度荷载耦合应力分析[J].郑州大学学报(工学版),2009,30(3):82-85.
- [4] 刘根昌.振捣式贫混凝土道路基层施工中裂缝的成因及防止[J].公路交通技术,2003(5):7-10.
- [5] 盛燕萍,陈拴发,郑木莲,等.免振捣多孔混凝土配合比设计方法[J].交通运输工程学报,2009,9(1):44-49.
- [6] SHENG Y P, LI H B, GUAN Bowen. Study on performance of normal and compaction-free porous concrete permeable base [J]. Advanced materials research, 2013, 345: 316-319.
- [7] SHENG Y P, LI H B, CHEN S F. Mix design of compaction-free porous concrete permeable base [J]. Advanced materials research, 2012, 368: 1416-1419.
- [8] 盛燕萍,陈拴发,李占全.免振捣多孔混凝土工作性研究[J].混凝土,2007(8):37-41.
- [9] 盛燕萍,陈拴发,郑木莲.不同成型方法的多孔混凝土路用性能对比[J].公路,2008(10):193-196.
- [10] 盛燕萍.免振捣多孔混凝土性能及其配合比设计方法研究[D].西安:长安大学公路学院,2006.
- [11] 孙丽娟.免振捣贫混凝土工作性研究[D].西安:长安大学公路学院,2010.
- [12] 林俊,姜小磊,姚佳良,等.路面基层碾压贫混凝土工作性影响因素试验研究[J].中外公路,2010,30(5):110-115.
- [13] 赵树林.免振捣贫混凝土性能及配合比设计方法研究[D].西安:长安大学公路学院,2007.
- [14] 梁思明,魏亚,张倩倩.路用碾压混凝土稠度与配合比研究[J].混凝土,2014(1):150-153.

(下转第 87 页)

改性凹凸棒石对水溶液中 Pb(II) 吸附性能研究

张 婕, 蔡 丽, 王景涛, 陈卫航

(郑州大学 化工与能源学院, 河南 郑州 450001)

摘 要: 以 3-氨丙基乙氧基硅烷(APTES)为改性剂,成功地合成改性凹凸棒石(ATP),并研究了其在水溶液中 Pb(II) 的吸附性能. 主要采取 SEM、XRD、BET、FTIR 和 TGA 对改性前后 ATP 的物理结构和化学组成进行测试分析;通过单因素静态实验,探讨了吸附时间、吸附剂用量、温度、Pb(II) 初始浓度等因素对改性 ATP 吸附 Pb(II) 性能的影响. 结果表明吸附时间 80 min, 吸附剂用量 $0.35 \text{ mg} \cdot \text{L}^{-1}$, 温度 $35 \text{ }^{\circ}\text{C}$, pH 为 6, 初始浓度 $100 \text{ mg} \cdot \text{L}^{-1}$ 时平衡吸附量为 $241.4 \text{ mg} \cdot \text{g}^{-1}$.

关键词: 凹凸棒石; 表面改性; 吸附性能; Pb(II)

中图分类号: O647.3

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.012

0 引言

冶金、电镀、化工、石油精炼等工业废水中含有铅、镉、汞等多种有害重金属离子. 其中铅主要经消化道及呼吸道被人体吸收,对造血系统、神经系统、消化系统造成损害. 治理废水污染的传统方法虽有很多,但吸附法因其高效、简便、易操作而被广泛采用.

凹凸棒石(ATP)是一种含有晶质水合镁、铝酸盐的粘土矿物^[1]. ATP 比表面积大、来源广、成本低、储量丰富,尤其在我国江苏盱眙县储量为 8.9 亿吨. ATP 在重金属离子吸附、有机污染物吸附等废水处理领域得到越来越广泛的研究和重视. 天然 ATP 是一种常见、廉价的吸附材料,但含有蒙脱土、石英、白英石及碳酸盐矿物等杂质,对 Pb(II) 的吸附量较低. 笔者选用 ATP,除对其常规提纯处理;还首次在表面接枝 APTES,对 Pb(II) 的吸附效果显著. 此改性过程为一步改性,过程简单,影响因素较少.

1 实验原理及方法

1.1 实验试剂及仪器

凹凸棒石(工业级,江苏省淮源矿业有限公司);盐酸(分析纯,烟台市双双化工有限公司);3-氨丙基乙氧基硅烷 APTES(分析纯,南京道宁化

工有限公司);硝酸(分析纯,洛阳市化学试剂厂);硝酸铅(分析纯,天津市科密欧化学试剂有限公司).

TAS-986F 原子吸收光谱仪(北京普析通用仪器公司).

1.2 ATP 的提纯

取适量 ATP 经溶解,高温煅烧^[2], 1 mol/L 盐酸, $80 \text{ }^{\circ}\text{C}$ 活化约 180 min 后,用蒸馏水洗涤至 pH 值约为 6, $30 \text{ }^{\circ}\text{C}$ 真空干燥,用粒径为 $\phi 178 \text{ mm}$ 的分子筛进行筛分.

1.3 ATP 的改性

在 50 mL 环己烷中加入提纯 ATP,该悬浮液在常温下超声 10 min 后,在快速搅拌的条件下按 $m(\text{APTES}):m(\text{ATP})=1:1$ 的比例加入 APTES, $60 \text{ }^{\circ}\text{C}$ 回流 20 h,静置后用无水乙醇离心洗涤, $60 \text{ }^{\circ}\text{C}$ 真空干燥.

1.4 样品的表征

SEM 采用日本电子株式会社生产的 JSM-7500F 型扫描电子显微镜.

XRD 采用日本理学公司的 D/MAX-RA 型 X 射线粉末衍射分析仪.

比表面积测试(BET)采用 GS-400 型气相层析仪,孔分布测定采用经典 N_2 吸附-脱附等温线,测试条件: $350 \text{ }^{\circ}\text{C}$, 4 h, $133.3 \times 10^{-5} \text{ Pa}$.

红外光谱(FTIR)采用 Thermo Nicolet IR200

收稿日期: 2017-06-16; 修订日期: 2017-08-27

基金项目: 国家自然科学基金资助项目(21206151)

作者简介: 张婕(1970—),女,河南郑州人,郑州大学副教授,博士,主要从事化学分离工程研究, E-mail: zhanglianbi@zzu.edu.cn.

型傅立叶红外光谱仪对 ATP 进行红外光谱分析.

热重分析(TGA)采用 TG-DTA 对改性前后 ATP 进行热重分析,升温速率为 10 °C/min.

1.5 吸附实验

取 20 ml 的 Pb(II) 溶液,加入吸附剂,恒温振荡离心,取上层清液用原子吸收分光光度计测 Pb(II) 浓度. ATP 对 Pb(II) 的单位吸附量和去除率可按公式(1)和(2)计算.

$$q_e = \frac{(C_0 - C_e) \times V}{m}; \quad (1)$$

$$R = \frac{C_0 - C_e}{C_0} \times 100\%, \quad (2)$$

式中: C_0 为初始浓度 $\text{mg} \cdot \text{L}^{-1}$; C_e 为平衡浓度, $\text{mg} \cdot \text{L}^{-1}$; V 为溶液的体积, L; m 为吸附剂的质量, g; q_e 为平衡吸附量, $\text{mg} \cdot \text{g}^{-1}$; R 为去除率, %.

对吸附剂用量考察时,取浓度均为 100 $\text{mg} \cdot \text{L}^{-1}$ 的 Pb(II) 溶液, pH 值为 3.5, 吸附剂用量在 0.25 ~ 1.5 $\text{mg} \cdot \text{L}^{-1}$, 25 °C 恒温振荡 180 min. 对平衡浓度和吸附温度考察时,初始浓度为 10 ~ 500 $\text{mg} \cdot \text{L}^{-1}$, pH 值为 3.5, 吸附剂用量为 0.35 $\text{mg} \cdot \text{L}^{-1}$, 在 25、35、45 °C 振荡 180 min. 对 pH 值考察时,初始浓度为 100 $\text{mg} \cdot \text{L}^{-1}$, pH 值为 1 ~ 6, 吸附剂用量 0.35 $\text{mg} \cdot \text{L}^{-1}$, 25 °C 恒温振荡 180 min. 对吸附时间考察时,初始浓度 100 $\text{mg} \cdot \text{L}^{-1}$, pH 值为 3.5, 吸附剂用量 0.35 $\text{mg} \cdot \text{L}^{-1}$, 25 °C 下振荡 0 ~ 300 min.

2 结果与讨论

2.1 ATP 提纯

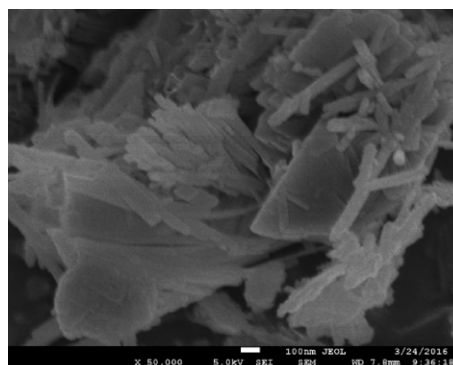
2.1.1 SEM 分析

提纯前后 ATP 扫描电镜(SEM)分析如图 1 所示.从图 1 可以看出,天然 ATP 表面杂乱无章且相互聚集,块状结构较多,针棒状较少.提纯后 ATP 整体分布变得疏松,晶体间矿物的连接物明显减少,针棒状较多.这表明高温煅烧和酸处理能够去除 ATP 表面杂质,去除碳酸盐,ATP 中八面体阳离子和少量四面体溶解,可导致 ATP 孔道疏通和孔径变大,比表面积增大,活性位点增加^[3].

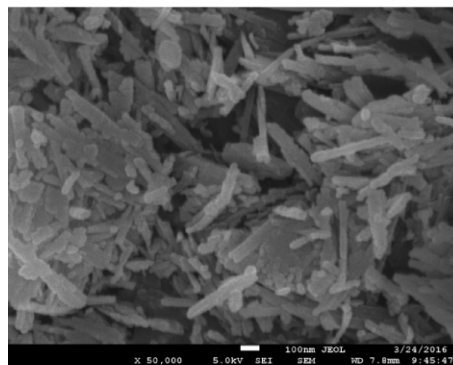
2.1.2 BET 分析

对提纯 ATP 进行 BET 测试,由图 2 看出, N_2 吸附-脱附等温线在低压 ($P/P_0 < 0.45$) 时, N_2 的吸附量缓慢增加,说明是单分子层吸附^[2-3]; 当压力相当高时 ($P/P_0 > 0.9$) 吸附量又急剧增加,出现毛细凝结现象,形成滞留回环,这发生在中孔吸附剂^[4].天然 ATP 比表面积约 9.6 ~ 36 $\text{m}^2 \cdot \text{g}^{-1}$, 通过提纯 ATP 比表面积增大为 170.58 $\text{m}^2 \cdot \text{g}^{-1}$,

说明提纯处理使其比表面积得到大幅度的增加,能提高 ATP 的吸附性能.



(a) 天然 ATP



(b) 提纯 ATP

图 1 天然 ATP (a) 和提纯 ATP (b) 的扫描电镜
Fig.1 SEM image of raw ATP (a) and purified ATP (b)

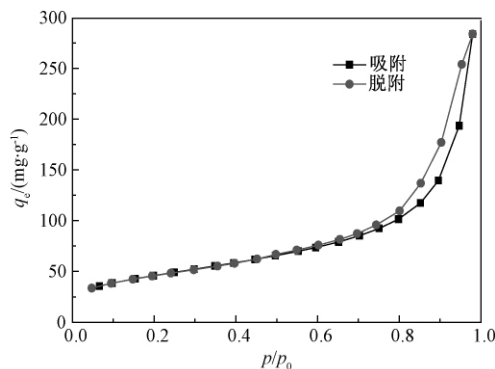


图 2 提纯 ATP 的 N_2 吸附-脱附等温线

Fig.2 N_2 adsorption-desorption isotherm of purified ATP

2.1.3 XRD 分析

天然 ATP 和提纯 ATP 进行 XRD 分析如图 2 所示, ATP 主要特征衍射峰的位置 ($2\theta = 8.34^\circ$ 、 27.5° 和 35.22°) 在提纯前后没有明显变化. 提纯 ATP 的石英 ($2\theta = 19.8^\circ$ 、 26.58° 和 40.1°), 白英石 ($2\theta = 29.06^\circ$) 和蒙脱土 ($2\theta = 16.02^\circ$)^[4] 特征衍射峰消失, 并且衍射峰 ($2\theta = 13.56^\circ$ 和 20.69°) 强度增强, 这是因为天然 ATP 表面杂质太多, 使

其表面特征衍射峰被掩盖,通过提纯处理可使其显露出来。

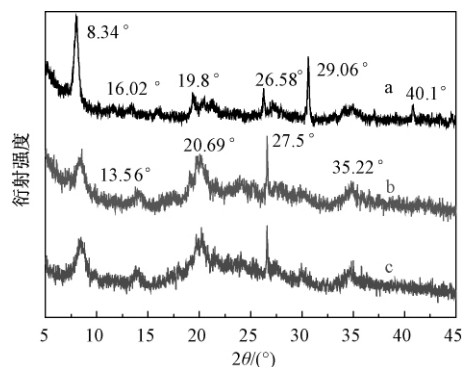


图3 天然 ATP(a), 提纯 ATP(b) 和改性 ATP(c) 的 XRD 谱图

Fig. 3 XRD patterns of raw ATP (a), purified of ATP (b) and modified of ATP (c)

2.2 ATP 的改性

2.2.1 FT-IR 分析

提纯 ATP 和改性 ATP 红外光谱分析如图 4 所示,与提纯 ATP 红外谱图相比,改性 ATP 红外谱图在 3442.7 cm^{-1} 、 1640.35 cm^{-1} 处的吸收峰强度明显减弱,这说明 APTES 与 ATP 表面的结合水接触,发生脱水缩合,形成 Si—O—Si,实现接枝。在 2923.46 cm^{-1} 出现新的吸收峰^[5],这是 APTES 上—CH₂—的不对称伸缩振动峰,说明 ATP 上的一OH 基团成功地与 APTES 发生了反应。

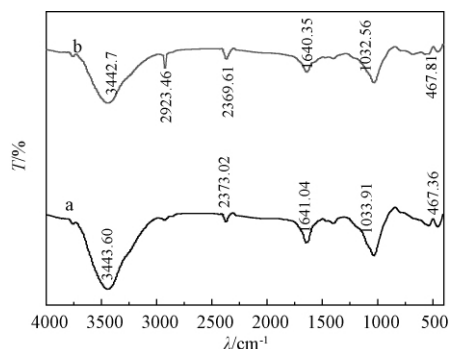


图4 提纯 ATP(a) 和改性 ATP(b) 的红外光谱
Fig. 4 FT-IR patterns of purified ATP (a) and modified ATP(b)

2.2.2 TGA 分析

对提纯 ATP 和改性 ATP 进行热失重分析,结果如图 5 所示。在热重分析前,样品都经过了高温 $400\text{ }^{\circ}\text{C}$ 煅烧,所以提纯 ATP 的失重率很低,在 $700\text{ }^{\circ}\text{C}$ 达到失重平衡,失重率仅为 4.5645% 。但是改性 ATP 失重现象非常明显,这是因为改性 ATP 表面接枝高分子量 APTES,在高温下有机物分解^[6],失重率增大,在达到失重平衡时,失重率

达 23.17% ,说明 APTES 成功地接枝 ATP 表面。

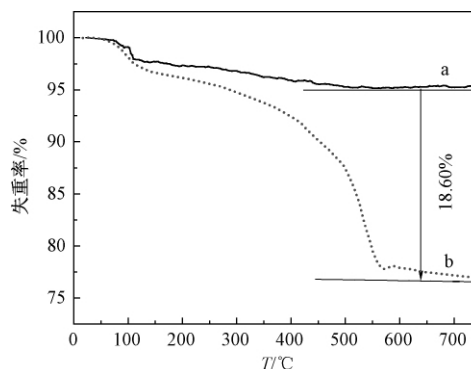


图5 提纯 ATP(a) 和改性 ATP(b) 的热失重谱图

Fig. 5 TGA patterns of purified ATP (a) and modified ATP (b)

2.3 改性 ATP 对 Pb(II) 吸附性能测试

在适宜的同等条件下比较改性前后的吸附量如表 1 所示。由表 1 可以看出,改性 ATP 吸附量是未改性 ATP 吸附量的 3.8 倍,这说明改性效果较为明显,同时也可以应用于其他离子的吸附。从表 2 可以看出,改性 ATP 吸附性能明显优于其他传统吸附剂,这对废水处理有很好的发展前景。

表 1 改性前后吸附性能对比

Tab. 1 Comparison of adsorption capacities of purified ATP and modified ATP

离子	饱和吸附量 $q_e / (\text{mg} \cdot \text{g}^{-1})$		提高率/%
	提纯	改性	
Pb(II)	98.72	341.36	245.8
Cd(II)	76.43	173.57	127.1

表 2 改性凹凸棒与其他传统吸附剂吸附性能对比

Tab. 2 Comparison of adsorption capacities of modified ATP and other adsorbents

离子	饱和吸附量 $q_e / (\text{mg} \cdot \text{g}^{-1})$	吸附剂	采用方法
Pb(II)	263.16	改性 ATP	本文方法
	19.60	活性污泥	文献[5]
	70.70	胺基化微球	文献[6]

2.3.1 吸附剂用量的影响

吸附剂用量对吸附 Pb(II) 的影响如图 6 所示。可以看出,随着改性 ATP 用量增加,吸附量逐渐减小,去除率却逐渐增加。这是因为溶液中的 Pb(II) 浓度一定,改性 ATP 的用量较小,单位质量 ATP 就会被周围较多的 Pb(II) 包围,因此吸附量大。但吸附剂的用量过小,改性 ATP 表面的氨基官能团数目有限,与 Pb(II) 螯合作用的机会少,所以 Pb(II) 去除率较低。但是当改性 ATP 用量增加,可以提供更多的吸附位点,其表面可以与

Pb(II) 结合的官能团数目增多,但是溶液中 Pb(II) 浓度是一定的,因此吸附量降低,去除率增大,溶液中 Pb(II) 去除率最终达到平衡. 为了保证实验吸附量和去除率都达到一定的水平,取 $0.35 \text{ g} \cdot \text{L}^{-1}$ 吸附剂,此时吸附量为 $200 \text{ mg} \cdot \text{g}^{-1}$,去除率 R 为 85%.

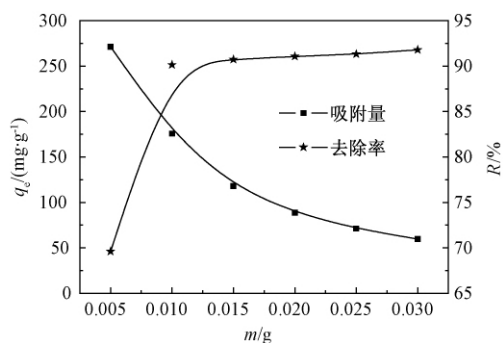


图 6 改性 ATP 用量的影响

Fig. 6 Effect of modified ATP dosage on adsorption

2.3.2 吸附时间的影响

吸附时间对改性 ATP 吸附的影响如图 7 所示. 可以看出,改性 ATP 对 Pb(II) 吸附初期是快速吸附过程,50 min 之后吸附速率明显减慢,直至 80 min 趋于吸附平衡. 这是因为初始阶段溶液浓度大,传质推动力就大,吸附剂表面的空闲活性位点多,与 Pb(II) 接触的机率就大,所以吸附速率就快. 随后改性 ATP 表面的空闲活性位点减少,使得吸附速率下降.

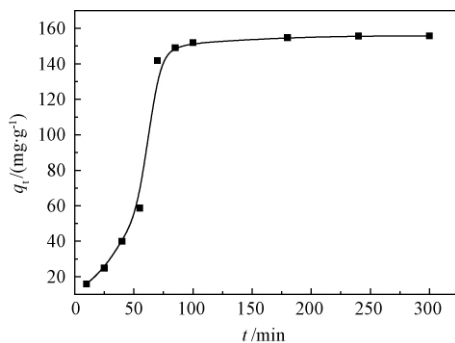


图 7 吸附时间的影响

Fig. 7 Effect of contact time on adsorption

2.3.3 初始浓度和温度的影响

初始浓度和温度对吸附 Pb(II) 的影响如图 8 所示. 可以看出,在同一温度下,改性 ATP 吸附量随着溶液中 Pb(II) 初始浓度增大而增大,这是因为 ATP 用量一定时,溶液中 Pb(II) 初始浓度越大,其与吸附剂的碰撞机会大,吸附量增大. 当溶液中的 Pb(II) 的初始浓度增大到一定值时,吸附量趋于平缓,这是因为改性 ATP 表面的氨基与 Pb(II) 发生螯合作用趋于饱和状态. 此外,改性

ATP 对 Pb(II) 的吸附量随着吸附温度的升高而增大,这说明该吸附过程是吸热过程,升高温度有利于吸附过程.

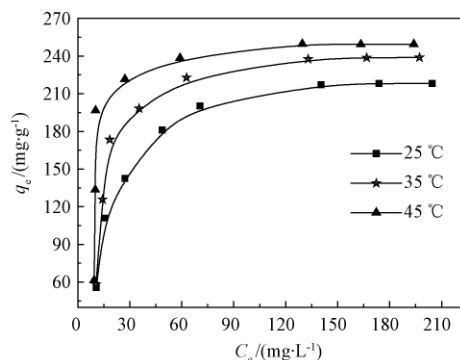


图 8 初始浓度和温度对吸附的影响

Fig. 8 Effect of initial concentration and temperature on adsorption

2.3.4 溶液 pH 值的影响

溶液 pH 值能够改变溶液中 Pb(II) 的存在状态和数量,故选用 pH 值范围为 1~6. 从图 9 看出:随着 pH 值的增大,吸附量逐渐增大,在 pH 值为 6 时达到最大吸附量 $329.57 \text{ mg} \cdot \text{g}^{-1}$;在 pH 值为 1 时,吸附量较低. 这是因为当溶液的 pH 值过低,改性 ATP 表面接枝的氨基基团上的氮原子被高度质子化,无法再与 Pb(II) 发生螯合作用,使吸附量偏低.

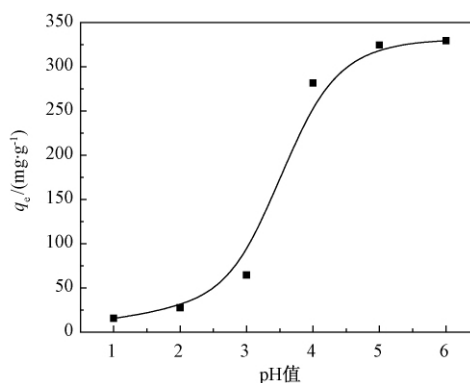


图 9 溶液 pH 的影响

Fig. 9 Effect of pH on adsorption

3 结论

笔者主要研究了 ATP 的提纯、改性和吸附性能三部分. 主要由 XRD、SEM、BET、FTIR、TGA 对天然 ATP,提纯 ATP 和改性 ATP 进行表征. 结果表明,提纯 ATP 的杂质明显减少,比表面积增大,成功地在 ATP 表面接枝上大量的氨基,对 Pb(II) 具有良好的吸附性能,吸附能力明显高于未改性 ATP.

参考文献:

- [1] 张彦灼,任琚,陶玲,等. 凹凸棒石粘土的物化性质研究进展[J]. 中国非金属矿工业导刊, 2013(1): 24-26.
- [2] 王秀芳,田勇,张会平. 高比表面积煤质活性炭的制备与活化机理[J]. 化工学报, 2009, 60(3): 733-737.
- [3] 何余生,李忠,奚红霞,等. 气固吸附等温线的研究进展[J]. 离子交换与吸附, 2004, 20(4): 376-384.
- [4] MANGUN C L, DEBARR J A, ECONOM Y J. Adsorption of sulfur dioxide on ammonia-treated activated carbon fiber[J]. Carbon, 2001, 39(11): 1689-1696.
- [5] MIRZABE G H, KESHTKAR A R. Selective sorption of U(VI) from aqueous solutions using a novel aminated $\text{Fe}_3\text{O}_4/\text{SiO}_2/\text{PVA}$ nanofiber adsorbent prepared by electro-spinning method[J]. J Radioanal Nucl Chem, 2015, 303(1): 561-576.
- [6] 张婕,尚成江,陈卫航,等. 胺基化杂化微球对溶液中 Cu 和 Pb(II) 的吸附[J]. 郑州大学学报(工学版), 2015, 36(5): 1-5.

Study on Absorption of Pb(II) from Aqueous Solution by Modified ATP

ZHANG Jie, CAI Li, WANG Jingtao, CHEN Weihang

(School of Chemical Engineering and Energy, Zhengzhou University, Zhengzhou 450001, China)

Abstract: With APTES as a modifier, the modified ATP was successfully prepared. And the adsorption of Pb(II) from aqueous solution by modified ATP was studied. The physical structure and chemical composition of raw ATP, purified ATP and modified ATP were characterized by SEM, XRD, BET, FTIR and TGA, respectively. The effects of contact time, the dosage of adsorbent, temperature and initial concentration on the adsorption property for Pb(II) were investigated through single factor static experiments. The results showed that the equilibrium adsorption capacity of Pb(II) was $241.4 \text{ mg} \cdot \text{g}^{-1}$ under the following conditions: contact time 80 min, the dosage of adsorbent $0.35 \text{ g} \cdot \text{L}^{-1}$, temperature 35°C , pH 6 and initial concentration $100 \text{ mg} \cdot \text{L}^{-1}$.

Key words: attapulgite(ATP); modified; adsorption; Pb(II)

(上接第82页)

Research on Workability of Compaction-free Lean Concrete Base

SHENG Yanping¹, SUN Shiwei¹, ZHAO Shulin¹, ZHOU Huili², CHEN Shuanfa¹

(1. School of Materials Science and Engineering, Chang'an University, Xi'an 710064, China; 2. Qinghai Research Institute Transportation, Xining 810008, China)

Abstract: A kind of compaction-free lean concrete (CFLC) base was put forward in this paper. The smoothness and stability of the base could be ensured by using compaction-free lean concrete base. Meanwhile, the construction machinery and construction cycle could be reduced. Also, construction environment could be improved. Factors influence workability of CFLC were analyzed, the method and index of evaluating workability of CFLC were advanced. Orthogonality test design was adopted in tests, regression formula of evaluating index-Vebe consistency value was deduced. The results showed that the workability of CFLC could be evaluated by Vebe consistency value.

Key words: road engineering; lean concrete; compaction-free; Vebe consistency value; workability

文章编号: 1671-6833(2017)06-0088-04

典型 MoSi_2 发热元件物相与显微结构分析

黄文江, 周 颖, 李 斯, 马成良, 史幸福, 王世界

(郑州大学 材料科学与工程学院, 河南 郑州 450001)

摘 要: 采用 XRD、SEM 和 EDS 等手段分别对 Kanthal 公司 Super 1700、Super 1800 和国产某厂三种不同 MoSi_2 发热元件的物相组成、微观形貌和微区元素成分进行分析, 探讨了非 MoSi_2 相对典型发热元件使用性能的影响, 结果表明: ①单质 Mo 对 MoSi_2 发热元件使用温度的提高非常不利; 在高温条件, 残留的单质 Mo 会与氧气反应生成气相的 MoO_3 , 使发热元件表面鼓泡并造成玻璃膜破坏. ②国产发热元件生产过程中应适当提高膨润土的加入量, 并严格控制膨润土中 K、Na 等活泼碱金属氧化物的量. ③通过添加 W 元素合金化的方法来制备复合 MoSi_2 粉并取代传统的 MoSi_2 粉, 可以提高发热元件的使用温度和使用寿命.

关键词: 发热元件; MoSi_2 ; 膨润土; 保护膜; 合金化

中图分类号: TB34 **文献标志码:** A **doi:** 10.13705/j.issn.1671-6833.2017.06.013

0 引言

MoSi_2 是一种金属间化合物, 具有陶瓷和金属的双重特性: 既具有类似金属的导电性, 又具有类似陶瓷的耐高温抗氧化性^[1]; 是一种潜在的可用于 1 500 °C 的高温结构材料. 目前 MoSi_2 材料最广泛的应用是作为高温电炉使用的发热元件^[2].

自从 1956 年瑞典 Kanthal 公司第一次将二硅化钼发热元件投放市场以来, 历经半个多世纪的发展, Kanthal Super 的系列产品一直引领着二硅化钼发热元件技术的进步, 最高使用温度从最初的 1 650 °C 提高到现在的 1 850 °C, 其产品在国际上长期处于垄断地位^[3]. 目前国内虽然已经有数十家单位能生产普通 MoSi_2 发热元件, 但其使用温度、使用寿命和室温力学性能等与 Kanthal 公司产品还有较大差距. 并且 Kanthal Super 元件在中性、还原性及碳化气氛下均能够顺利使用并产生良好效果, 这是国产发热元件所达不到的. 冯培忠等人^[4]研究了日本理研生产的发热元件与国产某元件的组成和结构, 提出发热元件的主要相包括灰色相(主要成分是 MoSi_2)和黑色相(主要铝硅酸盐玻璃相), 但是文章并没有指出两种

不同相的比例以及所含元素的种类及含量. 赵世柯等人^[5]研究了 Na_2O 对发热元件的特性的影响, 提出要降低发热元件中硅酸盐玻璃相中 Na 的含量. 艾云龙等人^[6]研究了 $\text{WSi}_2/\text{MoSi}_2$ 复合发热元件的制备及组织性能, 提出合金元素 W 的加入可以提高发热元件的抗弯强度和显微硬度, 提高烧损温度、降低热膨胀系数改善发热元件的热稳定性.

笔者利用 XRD、SEM、EDS 等分析手段, 对 Kanthal Super 系列与国产某厂发热元件的物相组成、显微组织和成分进行分析, 希望通过对对比分析出三种典型发热元件的相组成、成分和显微结构不同之处, 及其与发热元件使用性能的关系, 对我国的 MoSi_2 发热元件产品的性能提升提供有益参考.

1 试样制备和试验方法

试样采用瑞典 Kanthal 公司生产的 Kanthal-Super 1 700 和 Kanthal Super 1 800 发热元件, 它们在空气介质中的最高使用温度分别为 1 700 °C 和 1 800 °C^[7]; 河南登封某厂生产的发热元件, 在空气介质中最高使用温度大约为 1 650 °C. 采用 Philip PW-1700 型 X 射线衍射仪和 JEOL JSM-

收稿日期: 2017-05-15; 修订日期: 2017-07-28

基金项目: 国家自然科学基金资助项目(51172210, 51472220)

通信作者: 周颖(1968—), 女, 河南郑州人, 郑州大学副教授, 主要从事功能陶瓷研究, E-mail: zhouying@zzu.edu.cn.

5610LV 型扫描电镜(SEM) 配合 Feature Max 型能谱仪(EDS) 对国产及 KANTHAL 系列 MoSi_2 发热元件进行分析.

2 实验结果及分析

2.1 Mo_5Si_3 与 Mo 的影响

国产某厂、Kanthal Super 1 700、Kanthal Super 1800 发热元件的 XRD 谱图 1 所示. 从图 1 可以看出: 主要组成物相均为 MoSi_2 , 同时还均有痕量的 Mo_5Si_3 相存在. 文献 [8] 指出: Mo_5Si_3 抗氧化能力比较低, 大量存在会大大降低发热元件的高温抗氧化性能, 但是少量 Mo_5Si_3 的存在对其高温氧化性能没有太大影响, 并且适量的 Mo_5Si_3 存在可以使 MoSi_2 材料的高温强度得到比较明显的改善, 并有利于发热元件的高温使用. 此外, 图 1 显示出国产某厂生产的 MoSi_2 发热元件中还含有痕量单质 Mo 存在, 文献 [9] 指出, 发热元件在使用过程中, 残留的单质 Mo 在高温条件下会与氧气反应生成气态 MoO_3 , 气态 MoO_3 挥发会造成表面冒泡、玻璃态保护膜的破坏, 进而导致内部基体的暴露, 这可能是导致国产某厂生产的发热元件使用寿命缩短的原因之一. 因此, 在 MoSi_2 发热元件生产过程中要严格控制 Mo_5Si_3 的生成量, 同时杜绝单质 Mo 的存在.

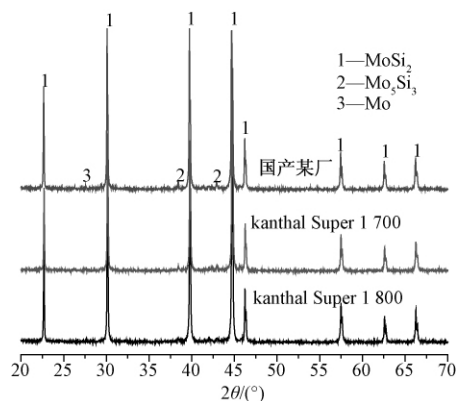


图 1 国产、Kanthal Super 1 700、Kanthal Super 1 800 发热元件的 XRD 图谱

Fig. 1 X-ray diffraction patterns for samples with three different kinds heating elements

2.2 铝硅酸盐玻璃相的影响

图 2 为国产某厂和 Kanthal 公司生产的 MoSi_2 发热元件的背散射 SEM 照片. 从图 2 中可以看出: 3 种不同 MoSi_2 发热元件主要由灰色基体相、黑色分散相和白色相 3 种物相组成. 结合灰色基体相、黑色分散相能谱数据(如表 1 所示) 和 XRD 结果, 可以确认灰色基体为 MoSi_2 ; 黑色相是一种以 Si 和 Al 为主的并包含有 K、Na、Mg 和 Ca 的物

质. 根据元件的组成可以认为黑色相是在制备过程中添加进去的铝硅酸盐材料, 但经过高温烧结, 铝硅酸盐材料已转变成玻璃态物质, 所以样品 XRD 结果也没有出现铝硅酸盐的峰, 因此可以确定黑色相是铝硅酸盐玻璃相. 结合文献 [10] 和 XRD 分析, 白色相是 Mo_5Si_3 , 它可能是在烧结过程中 MoSi_2 和坯体中残余的氧发生反应所致, 也可能是自蔓延合成 MoSi_2 粉体的副产物.

3 种典型 MoSi_2 发热元件在相分布均匀性及含量方面均存在明显的差别. 国产发热元件(如图 2(a) 所示) 灰色基体粒径较大且聚集严重, Kanthal MoSi_2 发热元件(如图 2(b) 、2(c) 所示) 灰色基体和黑色相分布比较均匀, 并且黑色相相对较多. 为了进一步弄清楚发热元件中黑色玻璃相所占体积百分比, 使用 Image-Pro Plus 6.0 软件对这 3 种 MoSi_2 发热元件黑色相面积进行了统计分析. 结果表明, 国产发热元件中黑色相所占比例最小(约为 3.49%), Kanthal 1 800 发热元件中黑色相所占比例最大(约为 8.58%); Kanthal 1 700 发热元件中黑色相所占比例居中(约为 6.66%).

目前生产二硅化钼发热元件的主要原料是二硅化钼粉和天然铝硅酸盐材料如膨润土. 膨润土的添加是为了提高发热元件的塑性成型能力, 同时促进 MoSi_2 的烧结, 并提高制品的致密性. 这是因为在高温烧结过程中, 添加的膨润土会熔化形成液相, 烧结时若有适当的液相出现, 往往会大大促进固相颗粒重排和传质过程, 这有利于烧结进行和制品的致密化. 当出现的液相量过多时, 不能保证产品的形状和尺寸, 引起制品变形, 并会降低发热元件的使用温度; 当液相量过少时, 残留气体不容易排出, 同样不利于发热元件的致密化. 因此, 要想获得致密度较高、气孔率较小的发热元件, 添加的膨润土的量一定要合适.

目前国产某厂 MoSi_2 发热元件致密度为 92.59%, 气孔率为 7.4%; 而进口发热元件的气孔率 < 1%, 致密度 98.6%, 国产发热元件的额定温度一般为 1 650 °C, Kanthal Super 1 700 型发热元件的额定温度为 1 700 °C [11]. 结合 Kanthal Super 1 700 和 1 800 型发热元件黑色相含量, 可以推测, 发热元件中黑色相含量与制品气孔率和致密度有一定相关性, 进而影响制品的使用额定温度和使用寿命. 因此, 在 MoSi_2 发热元件生产过程中要适当调整膨润土的加入量, 以得到低气孔率、高致密度的制品.

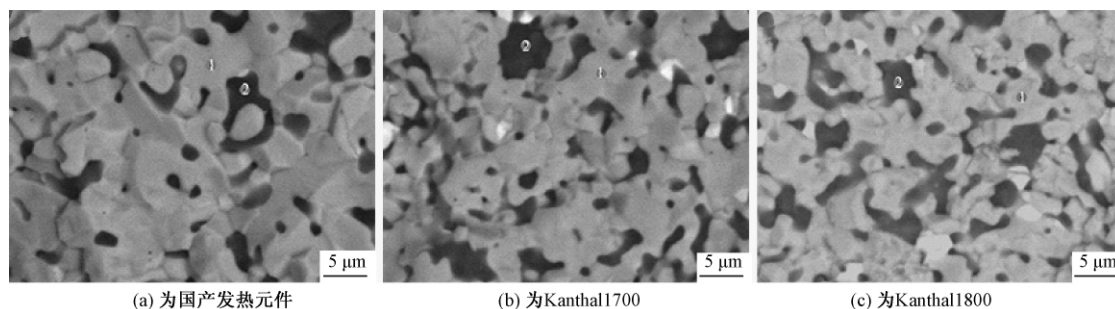


图2 发热元件的背散射 SEM 照片

Fig.2 Electron Backscatter Diffraction for samples with three different kinds heating elements

2.3 元素 W 和活泼金属元素 Na、K 的影响

表 1 为 3 种不同发热体灰色相和黑色相元素含量能谱分析数据. 从表 1 可以看出: 国产和 Kanthal 发热元件灰色基体所含主要元素是 Si 和 Mo, 但是不同元件的 Si 和 Mo(W) 的摩尔比例略有不同; 相比而言, Kanthal 发热元件的比值更接近 2, 此外 Kanthal 发热元件中有少量的 W 存在. 查阅文献可知: MoSi_2 的低温脆性和高温蠕变性大大影响了其实际运用, 合金化、复合化是改善这种性质的主要方法, 当少

量 Mo 元素被 W 部分取代制备的 $\text{WSi}_2/\text{MoSi}_2$ 复合发热元件时具有更好的高温强度, 并可以提高发热元件的使用温度^[12]; 并且 $\text{WSi}_2/\text{MoSi}_2$ 复合发热元件电阻率与 MoSi_2 发热元件相似, 随着温度的变化而变化, 没有老化现象, 不存在电阻不匹配的问题^[13]. 因此, 在发热元件生产制备过程中, 可以通过加入 W 或其他合金元素的方法得到复合 MoSi_2 粉体, 来取代传统的 MoSi_2 粉, 从而提高发热元件的使用温度和使用寿命.

表 1 发热元件灰色相和黑色相成分质量分数的 EDS 分析

Tab.1 Energy dispersive spectrum analysis for gray and black composition of heating element		%								
相组成	发热元件	Si	Mo	W	O	Al	Mg	Ca	Na	K
灰色相	国产某厂(a1)	64. 76	35. 40	0	0	0	0	0	0	0
	Kanthal Super 1 700(b1)	65. 56	31. 38	2. 97	0	0	0	0	0	0
	Kanthal Super 1 800(c1)	65. 92	20. 10	1. 93	0	0	0	0	0	0
黑色相	国产某厂(a2)	23. 80	0	0	68. 75	4. 62	0. 75	0. 48	1. 38	0. 22
	Kanthal Super 1 700(b2)	21. 60	0	0	66. 63	8. 73	1. 75	0. 94	0	0
	Kanthal Super 1 800(c2)	18. 38	0	0	70. 65	8. 40	1. 42	0. 34	0	0

发热元件生产过程中添加的膨润土的量及其性质是每个企业的核心机密之一, 因为这对发热元件的塑性成型性、烧结性和制品宏观性能等都有重要影响. 从表 1 可以看出: 3 种典型发热元件黑色相的元素组成相类似, 但元素含量有所不同. 其主要差别在于国产发热元件黑色相中含有少量的 Na 和 K 活泼碱金属元素存在; 而 Kanthal 发热元件黑色相中没有发现 Na、K 元素的存在, 或者是含量太少未能检测出来. Na、K 等活泼碱金属在膨润土中以 Na_2O 、 K_2O 的形式存在, 这些低熔点氧化物的存在会降低膨润土的熔化温度^[14], 改善铝硅酸盐熔体的高温流动性并降低熔体黏度, 这有利于熔体与 MoSi_2 颗粒接触并润湿和烧结致密性的提高. 但在发热元件高温使用过程中, 这些低熔点物质 Na、K 的挥发会产生“冒泡”现象, 并在发热元件表面形成白色斑点状区域, 这些区域

结构疏松并存在微小空洞, 严重损坏表面保护膜, 使发热元件的高温使用温度和使用寿命大大降低. 因此, 在 MoSi_2 发热元件制备过程中, 添加剂中 Na、K 等活泼碱金属氧化物的含量要严格控制.

3 结论

通过对国产某厂、Kanthal Super 1 700、Kanthal Super 1 800 3 种典型 MoSi_2 发热元件的物相、形貌和微区成分等对比分析, 得到如下结论:

(1) 典型 MoSi_2 发热元件的主要物相是 MoSi_2 相和少量的 Mo_5Si_3 相, 在制备 MoSi_2 粉体时适当控制钼硅比, 避免单质 Mo 的残留.

(2) 在 MoSi_2 发热元件生产过程中要适当调整膨润土的加入量, 根据本实验结果认为加入膨润土的重量百分比应控制在 8% 左右, 同时调整

膨润土的成分,严格控制 Na、K 等活泼碱金属氧化物的量。

(3) 通过加入 W 等合金元素的方法得到 Mo-Si₂ 复合粉体,来取代传统的 MoSi₂ 粉,以提高发热元件的使用温度和室温力学性能。

参考文献:

- [1] VASUDEVAN A K, PETROVIC J J. A comparative overview of molybdenum disilicide composites [J]. Materials science & engineering A, 1992, 155(1/2): 1-17.
- [2] 童桂英. 二硅化钼发热体力学性能的改进及其应用 [J]. 机械工程材料, 2003, 27(6): 49-49.
- [3] PERSSON P. Maximum element temperature for Kanthal Super 1800 s in flowing nitrogen atmosphere with low content of oxygen [EB/OL]. 2010. <http://www.diva-portal.org/smash/get/divaz:304269/FULLTEXT01>.
- [4] 冯培忠, 王晓虹, 缪姚军. 一种 MoSi₂ 发热元件的组成和结构分析 [J]. 硅酸盐通报, 2007, 26(5): 939-942.
- [5] 赵世柯, 王刚, 江莞. Na₂O 对 MoSi₂/Oxide 发热元件特性的影响 [J]. 材料科学与工艺, 2002, 10(2): 167-169.
- [6] 艾云龙, 程玉桂, 杨延清, 等. WSi₂/MoSi₂ 复合发热元件的制备及组织性能 [J]. 稀有金属材料与工程, 2005, 34(6): 962-965.
- [7] ZHANG C, SUO T, TAN W, et al. An experimental method for determination of dynamic mechanical behavior of materials at high temperatures [J]. International journal of impact engineering, 2017(102): 27-35.
- [8] SCHNEIBEL J H, SEKHAR J A. Microstructure and properties of MoSi₂-MoB and MoSi₂-Mo₅Si₃, molybdenum silicides [J]. Materials science & engineering A, 2003, 340(1/2): 204-211.
- [9] ZHANG Y Y, Li Y G, CHEN J, et al. Progress in Research on Mo-MoSi₂ functionally gradient material [J]. Advanced materials research, 2011, 28: 287-290: 636-641.
- [10] 魏猛. MoSi₂ 发热元件密度和使用温度的若干研究 [D]. 郑州: 郑州大学材料工程学院, 2011.
- [11] XU J, WANG Y, WENG B, et al. Preparation and characterization of MoSi₂/WSi₂ composites from MASHed powder [J]. Materials transactions, 2015, 56(3): 313-316.
- [12] 赵吉鹏, 邵松伟, 陈交贤. W 的添加对 MoSi₂ 显微组织的影响 [J]. 特种铸造及有色合金, 2011, 31(6): 504-506.
- [13] 艾云龙, 程玉桂, 杨延清, 等. WSi₂/MoSi₂ 复合发热元件的制备及组织性能 [J]. 稀有金属材料与工程, 2005, 34(6): 962-965.
- [14] 林锋. 硅钼棒表面“冒泡”现象的研究 [J]. 德州学院学报, 2005, 21(2): 87-89.

Study on Phase Composition and Microstructure of Typical MoSi₂ Heating Elements

HUANG Wenjiang, ZHOU Ying, LI Si, MA Chengliang, SHI Xingfu, WANG Shijie

(School of Materials Science and Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The phase composition, microstructure and micro-area elemental analysis of Kanthal and domestic MoSi₂ heating elements were studied by X-ray diffraction analysis(XRD), scanning electron microscope(SEM) and energy dispersive spectrometer(EDS). The impact of other substances except for the MoSi₂ phase on the performance of typical MoSi₂ heating elements was investigated. The results showed that: (1) Elemental Mo was detrimental to improving the service temperature of the MoSi₂ heating element. MoO₃ gas was formed by the reaction of residual Mo and oxygen at high temperature, resulting in the emergence of bubble on the surface and damage to the glass film. (2) It was necessary to increase the amount of bentonite during the production process of domestic heating element, while the amount of K, Na and other active alkali metal oxide in the bentonite was strictly controlled. (3) The MoSi₂ composite powders were synthesized via alloying method of adding W in replace of traditional MoSi₂ powders, which could improve the service life and operating temperature of the heating element.

Key words: heating element; MoSi₂; bentonite; protective film; alloying

文章编号: 1671-6833(2017)06-0092-05

$\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化歧化制备二甲基二氯硅烷的研究

徐文媛¹, 姚成姣¹, 徐永昌¹, 洪三国²

(1. 华东交通大学 理学院, 江西 南昌 330013; 2. 南昌大学 化学学院, 江西 南昌 330037)

摘 要: 一甲基三氯硅烷与三甲基氯硅烷的歧化反应可以制备最重要的单体二甲基二氯硅烷, 采用 B3LYP/6-311++G(3df,2pd) 方法对 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂体系进行了计算, 得到其主反应速控步的活化能为 $15.59\text{ kJ}\cdot\text{mol}^{-1}$. 实验表明: $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化制备二甲的最高产率为 81.30%, 比 $\gamma\text{-Al}_2\text{O}_3$ 的产率高, 结果与计算匹配.

关键词: $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$; DFT; 二甲基二氯硅烷; 歧化改性

中图分类号: O633

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2017.06.014

0 引言

Nature^[1]曾报道有机硅高性能材料^[2], 以无机结构为主链, 有机基团为侧链, 兼具无机、有机物的双重优点^[3], 而其重中之重在于单体二甲基二氯硅烷(简称二甲). 随着市场对二甲需求的日益增大, 直接合成法亦产生越来越多副产物, 如一甲基三氯硅烷(简称一甲)和三甲基氯硅烷(简称三甲). 因此如何将一甲和三甲歧化制备二甲这一变废为宝倍受关注. 德国拜耳公司^[4]研究出一种能够连续进行的低压歧化反应方法, 反应条件温和, 操作简单, 但催化剂用量很大, 而且还需要相当量的助剂, 成本高. 高压歧化法所需的压力一般为 3~10 MPa, 常用的催化剂是 AlCl_3 , 虽然二甲产率较高, 但该方法对设备要求高, 操作安全系数低. Kanner 等^[5]以 $\text{AlCl}_3\text{-CuCl}$ 为催化剂, 在 265 °C 及 2.52 MPa. 的条件下进行反应, 二甲含量可达 53.63%. 在成本低产量高的经济理念的号召下, 将副产物一甲和三甲采用常压歧化法来制备二甲. 不仅操作简单, 而且节约成本. 前期实验表明^[6]: $\gamma\text{-Al}_2\text{O}_3$ 和 $\text{NaAlCl}_4/\text{ZSM-5}$ 及 $\text{NaAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 均对歧化制备二甲有一定的催化活性. 而 Li 与 Na 属于同一主族, 具有相同的电子层排布. 因此推断 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 对歧化制备二甲也有一定的催化效果. 笔者采用 DFT^[7] 和 MP2^[8] 方法探讨了 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂的 L 酸^[9] 活性位-

AlCl 歧化一甲和三甲以制备二甲的反应历程并对其进行实验验证.

1 计算和实验方法

采用 Gaussian 03^[10] 程序中密度泛函 (B3LYP/6-311++G(3df,2pd))^[11] 方法对 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂催化过程中各驻点, 即反应物、中间物和产物的几何构型进行了全优化计算, 过渡态构型和反应路径则分别通过振动分析和 IRC 计算得到了确认^[12]. 最后利用二级微扰 MP2/6-311++G(3df,2pd) 方法对其进行了能量校正.

笔者采用高温高压负载法来制备 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂, 首先将 $\gamma\text{-Al}_2\text{O}_3$ 载体浸渍在一定浓度的 LiCl 溶液中, 浓度根据锂铝比进行调节, 然后一起放入 423.15 K 的烘箱中烘干, 再加入 AlCl_3 一起放入聚四氟乙烯棒中, 在马弗炉中 493.15 K 下加热 16 h. 待自然冷却后就制成催化剂.

2 结果与讨论

由 LiCl 与 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 制备 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂的能量变化图如图 1 所示, AlCl_3 负载后替换了 $\gamma\text{-Al}_2\text{O}_3$ 上原有的两个质子 H, 脱出 HCl, 形成了一 AlCl 的 L 酸活性中心和 LiCl 加在一起. 结果表明, 反应所需的活化能很低, 且产物的结构比反应物更稳定. 其中, R 表示反应物 LiCl

收稿日期: 2017-04-02; 修订日期: 2017-07-06

基金项目: 国家自然科学基金资助项目(21563011)

作者简介: 徐文媛(1975—), 女, 江西南昌人, 华东交通大学教授, 博士, 主要从事催化材料及反应机理研究, E-mail: xwyktz@163.com.

和 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$; TS 表示过渡态; P 表示产物 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$.

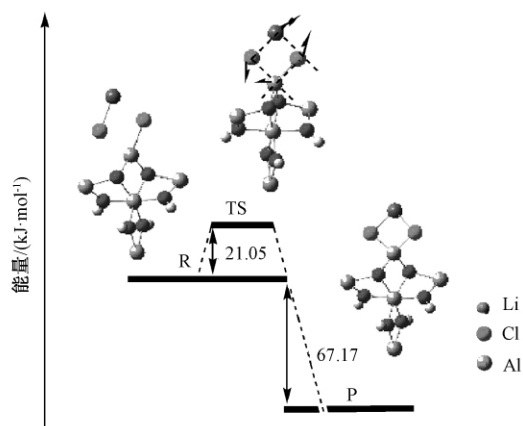


图1 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 生成反应的能线及过渡态走向产物的虚振模式图

Fig.1 Energy profiles in catalytic system of $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ and imaginary vibration mode of transition state

$\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂催化制备二甲反应

过程如图2所示,首先复盐 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 在反应起始先还原为 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 和 LiCl ,露出 L 酸活性位点 -AlCl ,再与反应体系开始作用. 反应分为两个通道: 通道一中三甲 ($(\text{CH}_3)_3\text{SiCl}$) 首先与催化剂 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 反应经过渡态 TS1,生成中间体 I ($\text{AlCH}_3\text{Cl}_2/\gamma\text{-Al}_2\text{O}_3$) 与二甲,接着一甲 (CH_3SiCl_3) 与 I 反应经过渡态 TS2, 还原出催化剂,并得所需产物二甲,此通道为反应的主通道; 通道二中一甲则先跟催化剂进行官能团互换,生成中间体 I ($\text{CH}_3\text{AlCl}_2/\gamma\text{-Al}_2\text{O}_3$) 和副产物四氯化硅,接着三甲与 I 反应,还原出催化剂 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 和副产物四甲基硅烷(简称四甲),最后 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 又可与 LiCl 结合成复盐,循环参与反应.

图3表示 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂参与反应关键原子间距沿 IRC 的变化. 从反应开始到结束关键原子间距都呈逐渐缩短或拉长的趋势,证实了过渡态与反应物、产物的正确连接,因此确认了计算所得过渡态是真实可靠的. 以 $\text{R1} \rightarrow \text{P1}$ 图为例,并对比图2中 TS1 的虚振模式图,三甲中

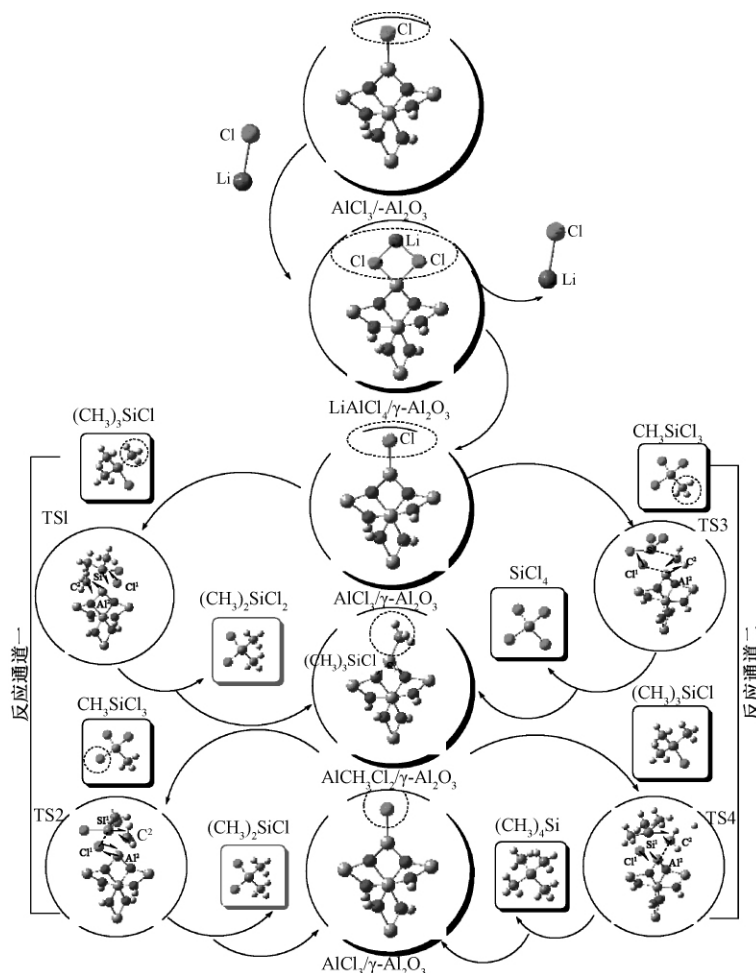


图2 关键原子及过渡态走向产物的虚振模式及反应流程

Fig.2 Key atomic number, imaginary vibration mode of transition state and reaction process

C^2-Si^1 和 $LiAlCl_4/\gamma-Al_2O_3$ 中 Cl^1-Al^2 键呈拉长趋势最后断开,而 Al^2-C^2 和 Si^1-Cl^1 逐渐靠拢,最后形成新键,其他的图类似分析,不再赘述。

图 4 是 $LiAlCl_4/\gamma-Al_2O_3$ 催化剂歧化制备二甲各步骤的反应热和活化能数据。由实验组前期研究可知^[13], $\gamma-Al_2O_3$ 催化剂歧化制备二甲的活化能为 $134.59 \text{ kJ}\cdot\text{mol}^{-1}$ 。对比图 4 可知,通道一

主反应的速控步为经 TS1 步,其活化能为 $15.59 \text{ kJ}\cdot\text{mol}^{-1}$;通道二副反应的速控步为经 TS3 步,其活化能为 $174.78 \text{ kJ}\cdot\text{mol}^{-1}$ 。可见,通道一的主反应比通道二的副反应容易发生;同时,与未负载 $AlCl_3$ 的 $\gamma-Al_2O_3$ 催化体系相比,一 $AlCl$ 引入使得反应明显容易进行,因此 $LiAlCl_4/\gamma-Al_2O_3$ 催化剂歧化制备二甲的催化效果较好。

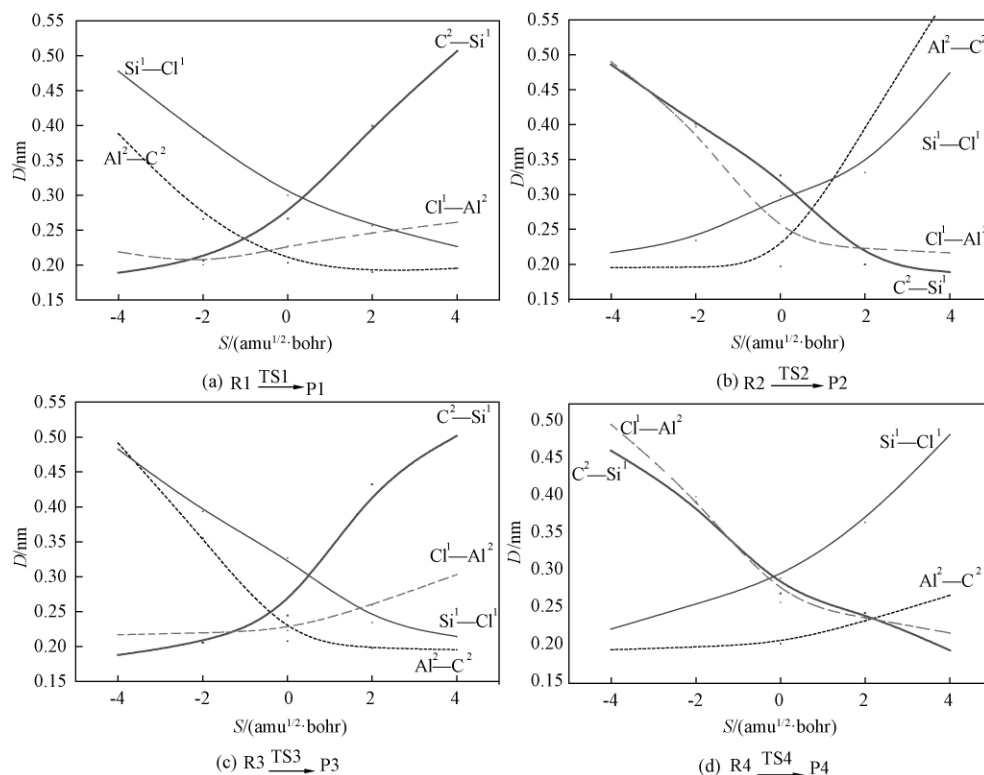


图 3 关键原子间距沿 IRC 的变化趋势

Fig. 3 The spacing trend of key atoms along IRC

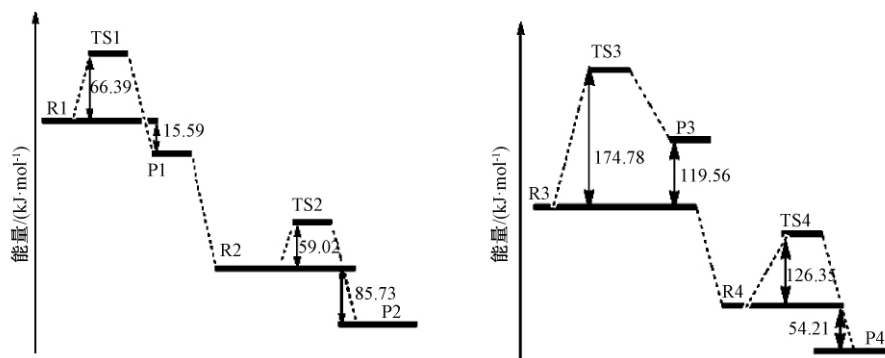


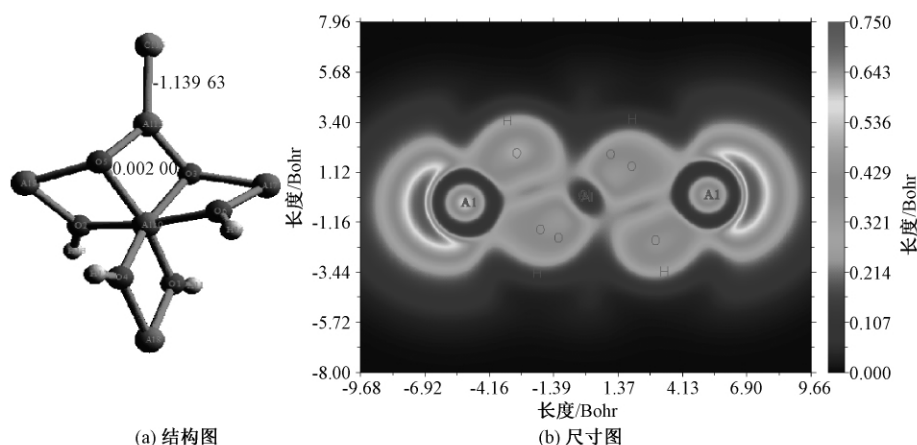
图 4 催化体系反应能线图

Fig. 4 Energy profiles in catalytic system

为了进一步探究催化剂的活性情况,结合 LOL 填色图对其进行键级分析^[14],计算得到 $Li-AlCl_4$ 脱去 $LiCl$ 后 $AlCl_3/\gamma-Al_2O_3$ 催化剂中 $Al^{15}-Cl^{16}$ 的 Mayer 键级为 $-1.139\ 63$, Al^{15} 、 Cl^{16} 、 O^3 、 O^5 的四中心键级为 $0.002\ 00$ 。为更直观地观察电子的定域性情况,对 $AlCl_3/\gamma-Al_2O_3$ 催化剂进行 LOL

分析,如图 5 所示。

由图 5 可知,中间的 Al 与 Cl 成键,两者同时位于离域性强区域,说明 $Al-Cl$ 键的电子能共享,同时表明催化剂 L 酸性的来源正出自于此。左右两侧的 Al 原子电子离域性居中,而其原子的周边电子离域能力较强,能与其他 Al 进行电子

图5 $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ 催化剂的 LOL 分析图Fig. 5 LOL analysis of $\text{AlCl}_3/\gamma\text{-Al}_2\text{O}_3$ catalyst

离域共享. 所有的 O 原子离域性居中, 这与 O 作为 $\gamma\text{-Al}_2\text{O}_3$ 的 Al 与 Al 间的结构氧桥没有明显活性是匹配的. 所有的 H 原子都在离域性较强的区域, 说明这些 H 可以作为质子的活性中心参与反应, 同时也可以与活性组分 LiAlCl_4 结合形成新的活性中心参与反应, 这与我们之前得到 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 的形成过程是吻合的. 为了进一步证实其催化活性, 对其进行了实验验证, 如图 6 所示.

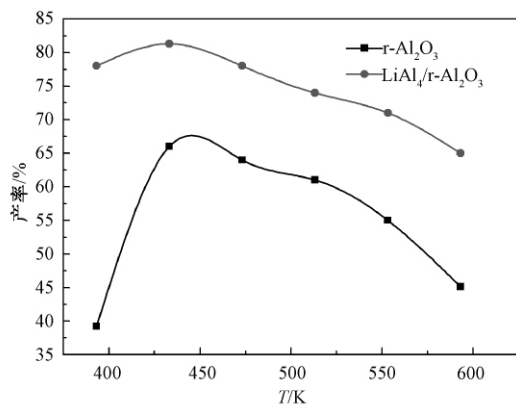


图6 温度对产率的影响

Fig. 6 Temperature affects Productivity

分别取 $\gamma\text{-Al}_2\text{O}_3$ 和 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 适量填充反应管, 置于反应釜中, 温度分别设定为 393.15、433.15、473.15、513.15、553.15、593.15 K. 反应物一甲和三甲的质量比为 1:1, 实验时取二者的混合物 4 μL 进样. 其他实验方面的证据见文献 [15]. 由图 6 可知, 反应温度在 433.15 K 下的产率最高为 81.30%, 说明此温度下催化活性最好. 对比 $\gamma\text{-Al}_2\text{O}_3$ 和 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 的实验数据, 并和前面的计算结果进行对比分析可知, $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂的催化活性较好. 在此基础上, 又对两种催化剂进行了红外表征, 如图 7 所示.

图 7 是 $\gamma\text{-Al}_2\text{O}_3$ 和 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 两种催化

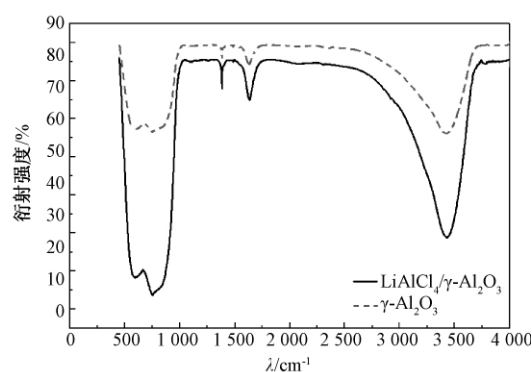


图7 催化剂的红外谱图

Fig. 7 FTIR of catalyst

剂的红外谱图, 在 1 450 cm^{-1} 附近处属于弱酸 L 酸性吸收峰, 在 1 540 cm^{-1} 处属于强酸 B 酸性吸收峰. 由图 7 可以看出, 催化剂 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 的 L 酸性和 B 酸性均强于载体 $\gamma\text{-Al}_2\text{O}_3$. 催化剂的酸性越强, 其催化性能越好. B 酸位为质子 H 活性中心, 而 L 酸位为负载 LiAlCl_4 后的 Al-Cl 键提供, 两者均为歧化的活性位. 实验结果和前述理论计算结果匹配.

3 结论

采用 DFT、MP2 和实验的方法对 $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化剂歧化一甲基三氯硅烷和三甲基氯硅烷以制备二甲基二氯硅烷的反应, 得到结论如下:

(1) $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化体系主反应速控步骤是第一步, 活化能为 15.59 $\text{kJ}\cdot\text{mol}^{-1}$, 催化效果比 $\gamma\text{-Al}_2\text{O}_3$ 好.

(2) 通过键级分析和 LOL 分析得 $\text{Al}^{15}\text{-Cl}^{16}$ 的 Mayer 键级为 -1.139 63, Al^{15} 、 Cl^{16} 、 O^3 、 O^5 的四中心键级为 0.002 00, 说明 Al-Cl 键易活化以提高反应的 L 酸中心.

(3) 由实验结果可知, 催化剂的酸性与活性正相关, $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化活性与酸性均比 $\gamma\text{-Al}_2\text{O}_3$ 好, 反应温度在 433.15 K 下的产率最高为 81.30%, 与计算结果匹配。

参考文献:

- [1] FREGERT S, RORSMAN H. Allergy to a carbon-functional organic silicon compound, dimethyldi-(4-hydroxyphenyl)-silane [J]. *Nature*, 1961(192): 989–900.
- [2] 佚名. 蓝星有机硅推出防火阻燃新产品 [J]. *有机硅材料*, 2014, 4(4): 266–267.
- [3] GORDON A D, HINCH B J, STRONGIN D R. Effects of individual promoters on the direct synthesis of methylchlorosilanes [J]. *Journal of catalysis*, 2009, 266(2): 291–298.
- [4] BOKERMAN G N, CANNADY J P, DGILVY A E. Production of organosilanes from polysilanes [P]. US: 5175329, 1992.
- [5] KANNER B, LEWIS K M. Process for preparing organosilanes utilizing copper halide-aluminium halide catalysts [P]. USP: 4599441, 1986–07–08.
- [6] 徐文媛, 王利伟. $\text{NaAlCl}_4/\text{ZSM-5}$ 催化甲基三氯硅烷歧化反应性能 [J]. *郑州大学学报(工学版)*, 2015, 36(5): 25–29.
- [7] BECKE A D. Density-functional thermochemistry. III: The role of exact exchange [J]. *The journal of chemical physics*, 1993, 98(7): 5648–5652.
- [8] SIMONS J, JORGENSEN P, HELGAKER T U. Higher molecular-deformation derivatives of the configuration interaction energy [J]. *Chem Phys Lett*, 1984, 86(3): 413–432.
- [9] XU W Y, LIU Y P, ZHOU J X, et al. Transforming bronsted acid to lewis acid on ZSM-5 disproportionation catalyst before and after loading AlCl_3 [J]. *Asian journal of chemistry*, 2015, 27(3): 1147–1152.
- [10] FRISCH M J, TRUCKS G W, SCHLEGEL H B, et al. Gaussian 03, Revision B. 04 [CP]. Pittsburgh, PA: Gaussian, Inc, 2003.
- [11] JASINSKI R, BARANSKI A. Molecular mechanism of Diels-Alder reaction between (E) 3-trichloro-1-nitropropene and cyclopentadiene: B3LYP/6-31G(d) computational study [J]. *Turkish journal of chemistry*, 2013, 37(5): 848–852.
- [12] BAO Y F, LIU Z Z, LIU D S, et al. Theoretical study on the structural and optoelectronic properties of the linear perfluorooctane sulfonate (PFOS) [J]. *Chinese J Struct Chem*, 2013, 32(9): 1384–1356.
- [13] 万欢欢. $\text{MeAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 系列催化剂制备二甲基二氯硅烷的机理研究 [D]. 南昌: 华东交通大学理学院, 2016.
- [14] LU T, CHEN F W. Multiwfn: A multifunctional wavefunction analyzer [J]. *Journal of computation chemistry*, 2012, 33(5): 580–592.
- [15] 刘一平. 氧化铝系列催化剂歧化甲基三氯硅烷反应的研究 [D]. 南昌: 华东交通大学理学院, 2015.

The Theoretical Research of Dimethyldichlorosilane Catalyzed by $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$

XU Wenyuan¹, YAO Chengjiao¹, XU Yongchang¹, HONG Sanguo²

(1. College of Science, East China of Jiaotong University, Nanchang 330013, China; 2. College of Science, Nanchang University, Nanchang 330037, China)

Abstract: Most important monomer dimethyldichlorosilane could be prepared by disproportionating methyltrichlorosilane and chlorotrimethylsilane. Disproportionation catalyzed by $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ were calculated at B3LYP/6-31++G(3df,2pd) level. Activation energies of the rate-determining step of main reaction was 15.59 kJ·mol⁻¹. The experiment indicated the productivity that $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ catalyst preparing dimethyldichlorosilane is 81.3%. It was higher than $\gamma\text{-Al}_2\text{O}_3$ catalyst. and it coincided with calculated results.

Key words: $\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$; DFT; dimethyldichlorosilane; disproportionation

二〇一七年总目次

土木与交通工程

蝉翼结构静力学仿真和力学性能分析	徐 赫, 杨 怡, 雷志鹏 (1-1)
基于动态特性的固定连接杆形状参数研究	曾发林, 曾少波, 李建康 (1-6)
交通荷载作用下低填道路软土地基的动力响应分析	张 浩, 杨 玲, 郭院成 (1-11)
土压平衡盾构土仓排土引起的干砂地层响应特征分析	瞿同明, 王树英, 刘朋飞 (1-16)
基于高铁站区影响的城市空间研究	曹 阳 (2-21)
高速列车隧道等速交会条件下人体舒适度分析	张运良, 张志成, 杨伟超, 丁铭鸿 (2-26)
基于集对分析的城市道路交通安全性评价	张劲楠, 钱晓东 (2-30)
基于 ARMA 和卡尔曼滤波的短时交通预测	杨高飞, 徐 睿, 秦 鸣, 郑凯俐, 张 兵 (2-36)
改进蚁群算法在交通分配模型中的应用	常玉林, 汪小淳, 张 鹏 (2-41)
信号交叉口禁左车流交通组织设计	李爱增, 贾俊波, 张亚飞, 李文权 (2-45)
冻融循环作用下橡胶自密实混凝土力学性能研究	张卫东, 王振波, 何卫忠 (2-78)
变截面波形钢腹板组合箱梁剪应力及剪力传递效率分析	武海鹏, 李 杰, 陈 淮 (2-83)
基于性能驱动的校园食堂建筑被动式节能设计	张建涛, 周丽华 (3-92)
风化花岗岩地层隧道衬砌裂损分析与处理研究	孙雁军, 阳军生, 张 箭, 王树英 (4-78)
悬挂式单轨交通线路适应性研究与分析	刘文龙, 徐延海 (4-84)
在役钢筋混凝土 T 形梁桥承载能力的模糊随机评估法	
.....	翟科玮, 杜朝伟, 郑凤玺, 郑元勋 (4-88)
关于凸极小化的 Douglas-Rachford 分裂方法的一个注	康倍倍, 董云达, 王亚丽 (4-94)
高速公路拓宽中桩类型优选与板桩适应性分析	李海滨, 柯胜旺, 申艳军 (5-44)

II

钢纤维混凝土路面接缝传荷性能衰减规律的试验研究	王建宁, 窦远明, 孙吉书, 魏 明, 翟玉玺 (5-50)
钢桥面铺装层粘接强度超声波无损检测技术研究	吕彭民, 吴玉文, 宋年波 (5-55)
大型预应力混凝土矩形渡槽槽身结构型式研究	陈玉英 (5-61)
基于响应面法的车内噪声分析与优化	陆森林, 任栢翔 (5-65)
路面免碾压贫混凝土基层工作性研究	盛燕萍, 孙仕伟, 赵澍林, 周辉丽, 陈拴发 (6-79)

化工与材料工程

五层粘接结构的断裂行为模拟研究	张 军, 李志宏, 贾 宏 (1-22)
火源横向位置对隧道火灾烟气分岔流动影响	钟 委, 端木维可, 李华琳, 梁天水 (1-27)
葡萄糖与二价锰离子对磺胺-5-甲氧基嘧啶钠生物降解影响	郭夏丽, 张凡英, 陈红坤, 罗亚婷 (1-32)
梭型凹凸板式换热器相变换热的优化研究	王定标, 张喜迎, 韩 勇, 邓 静, 董 桢 (1-36)
一种强化换热流道的传热特性及流体扩散性分析	王永庆, 朱 冰, 赵 頔, 靳遵龙 (1-41)
350 MW 超临界机组膜式水冷壁壁温计算	王为术, 崔 强, 郑梦星, 陈 刚 (1-46)
原位聚合 PA66/SiO ₂ 纳米复合材料制备及其性能研究	郑 瑾, 张振亚, 王 龙, 刘文涛, 朱诚身 (1-83)
p-n 型 CuO/ α -Fe ₂ O ₃ 复合半导体材料的制备与气敏性能研究	秦 聪, 王 燕, 陈泽华, 孙 广 (1-87)
聚乙烯醇分散对 LiFePO ₄ 薄膜气敏性的影响	帕提曼·尼扎木丁, 阿布力孜·伊米提 (1-92)
填料对再生沥青胶浆黏弹响应的影响	陈华鑫, 牛昌昌, 况栋梁, 田俊壮, 杨 凯 (2-72)
氧化铜掺杂氧化锌气敏材料的制备及气敏性能研究	宣天美, 孙健武, 葛美英, 尹桂林, 何丹农 (2-88)
rGO/ZnSn(OH) ₆ 复合材料的合成及其光催化性能研究	卢红霞, 李雪雪, 雷 君, 赵天歌, 邵 刚 (3-1)
Zn 对 Al-Cu-Mg-Ag 合金组织和常温力学性能的影响	王杰芳, 张 铎, 郭巧能, 刘忠侠, 王明星 (3-6)

内置扭带管内湍流流动与传热数值模拟	吴金星, 王超, 王明强, 刘艳会, 李亚飞 (3-10)
增塑剂改性沥青路用性能试验研究	傅珍, 申万青, 孔志峰, 张超 (3-15)
微细粒褐煤中腐植酸的水溶特性及其影响研究	邢龙杰, 黄艳芳, 王文娟, 韩桂洪, 刘炯天 (3-20)
改性荷叶对 1,4-苯二酚吸附性能的研究	韩秀丽, 宋阿娟, 白净, 方书起 (3-25)
Ta 掺杂对 $\text{BaCe}_{0.8}\text{In}_{0.1}\text{Y}_{0.1}\text{O}_{3-\delta}$ 电解质材料性能的影响 ...	苏金瑞, 南怡晨, 何豪, 赵莹莹 (3-30)
吸附剂浓度对 $\text{Cu(II)}/\text{HP-PAA}/\text{SiO}_2$ 吸附体系的影响 ...	范忠雷, 马翔宇, 李瑞瑞, 乔欢欢 (3-35)
基于 ANSYS CFX 的汽轮机叶片气动性能优化模拟研究	周俊杰, 王璞, 周金方 (4-1)
纵向激励下防波板组合形式对罐箱液体晃动的影响	丁昌, 付延唐, 吴学红, 龚毅 (4-7)
基于包埋功能微生物的生物滤塔净化甲苯废气研究	朱亚中, 李顺义, 骆翼梦, 马宏业, 王岩 (4-12)
颗粒直径与轴向分布对吸附热变换器传热传质的影响	盛遵荣, 薛冰, 刘周明, 魏新利 (4-17)
铈氧化物/石墨烯的制备及其对 PO_4^{3-} 的吸附性能	史春燕, 范冰冰, 李娅娅, 胡永宝, 张锐 (4-23)
功能离子液体催化 CO_2 与环氧化合物的加成反应	吴晶敏, 李伟杰 (5-71)
改性凹凸棒石对水溶液中 Pb(II) 吸附性能研究	张婕, 蔡丽, 王景涛, 陈卫航 (6-83)
典型 MoSi_2 发热元件物相与显微结构分析	黄文江, 周颖, 李斯, 马成良, 史幸福, 王世界 (6-88)
$\text{LiAlCl}_4/\gamma\text{-Al}_2\text{O}_3$ 催化歧化制备二甲基二氯硅烷的研究	徐文媛, 姚成姣, 徐永昌, 洪三国 (6-92)

电气工程

一种双余度无刷直流电机转矩脉动抑制控制策略研究	刘卫国, 霍达, 谭博, 公超 (1-50)
基于区间样本和回声状态网络的风电功率不确定性预测	孙晓燕, 时良振, 徐瑞东, 张勇 (1-56)

一种多特征提取及融合的虹膜识别方法	张 震, 刘 博, 李 龙 (1-63)
非均衡交通流五岔路口交通信号智能控制研究	董海龙, 汤旻安, 程海鹏 (1-68)
基于 K-Means 和时间匹配的位置预测模型	胡 燕, 朱晓瑛, 马 刚 (2-17)
基于 PMU 的降阶二次状态估计算法研究	蒋建东, 杜耀恒, 燕跃豪, 鲍 薇 (2-50)
基于双贝叶斯估计的动态威胁运动状态估计	刘 厂, 赵俊翔, 胡 海 (2-55)
基于降维数据边界点曲率的变电站设备识别 窦本君, 纪 勇, 郑尚高, 冯冬青, 罗 勇 (2-61)
解码动物转向行为的 ICA-小波特征提取方法	董芳芳, 尚志刚, 刘新玉, 万 红 (3-39)
输电固定成本分摊的 AR-DEA 合作博弈核仁解方法	王金凤, 孟徐飞, 王 铮, 杨丽徙 (3-44)
结合投影与近邻操作的支持向量快速筛选方法	李蒙蒙, 尚志刚, 李志辉 (3-49)
不同类型供电区域间典型供电模式过渡方案研究 杨 卓, 王利利, 杨 博, 关朝杰, 程子霞 (3-54)
真空开关低分散性永磁操动机构控制系统设计 程 显, 陈占清, 何 周, 袁端磊, 钟建英 (3-58)
一种基于聚类分析的自适应步态检测方法	姜 鸣, 赵红宇, 刘学良 (3-63)
基于 Mean Shift 聚类的多级阈值化方法	程 全, 刘晓青, 刘玉春, 王志良 (6-64)
石油钻井工程预警技术发展概述	李 广, 张 衡, 王 杰, 朱晓东, 岳彩通 (6-70)
具备直流故障隔离能力的新型 MMC 拓扑研究	张清枝, 李志辉 (6-74)

机械工程

全矢谱技术在 WTDS 实验台故障诊断中的应用	韩 捷, 张钱龙 (1-74)
基于螺栓联接非线性系统的杆式超声电机建模与分析	陶 征, 胡 斌, 刘 旭 (1-78)
双钢轮压路机的非线性建模及其传动干扰分析	沈培辉, 林述温 (2-66)
钻井力矩电机新型无位置传感器控制方法	白国长, 张宇昕 (3-68)
弹性底板上的液压支架整体尺寸参数优化	王 志, 王朝雅, 杨 飞 (3-73)
18CrNiMo7-6 高速磨削条件下表面质量的试验研究	王 栋, 刘昱范, 陈 鑫 (3-78)
基于滑移率的 CVT 夹紧力控制策略研究	袁晓红, 陈 龙, 汪怡平 (3-82)

基于位移响应最小化的约束阻尼结构优化设计	侯俊剑, 房占鹏, 何文斌 (3-87)
快递与第三方智能快递柜合作机制的演化博弈分析	李玉民, 杨 露, 王新露 (5-81)
带批处理机的多阶段柔性流水车间调度优化	王君妍, 王薛苑, 轩 华 (5-86)
飞碟游乐设备驱动轴疲劳失效分析	刘治华, 刘博见, 许伟超, 张银霞, 陶德岗 (5-91)
考虑公差原则的三维公差分析方法研究	赵凤霞, 金少搏, 李纪峰 (6-39)
计及线路损耗的峰谷分时电价优化模型	赵国生, 詹天乐, 李 博 (6-45)
基于 Möbius 变换的复有理圆弧样条	任两品, 薛均晓, 张朝阳, 王定标 (6-50)
基于 FPGA 的直流电机伺服控制系统设计与实现	常怡萍, 刘 琴, 杨 蕾 (6-54)
考虑顾客时间紧迫度的生鲜电商配送路径优化问题	
.....	崔 岩, 张子祥, 时 新, 王晓亮, 王振锋 (6-59)

信息工程

基于人工鱼群优化 SVM 的声磁标签信号检测研究	邓记才, 耿亚南 (4-35)
SINS 姿态估计的 HoMM-UKF 模型算法	丁国强, 张 铎, 熊 明, 周卫东 (4-39)
基于 PSO-RBF 神经网络的雾霾车牌识别算法研究	毛晓波, 张 群, 梁 静, 刘艳红 (4-46)
基于概率自适应蚁群算法的云任务调度方法	
.....	王俊英, 颜芬芬, 陈 鹏, 董方敏, 臧兆祥 (4-51)
基于图的概念重现发现与预测	白 洋, 王志海, 孙艳歌 (4-57)
大规模 MIMO 系统中功率分配算法的能效研究	陆彦辉, 苗盼盼, 杨守义 (4-65)
WSN 中基于离散人工鱼群的分簇拓扑优化算法	阎新芳, 张晓丹, 严晶晶, 冯 岩 (4-69)
基于 Voronoi 盲区的三维无线传感器网络覆盖优化算法	方 伟, 梅希薇 (4-73)
基于互信息和 Just-in-Time 优化的回声状态网络	张 衡, 王河山 (5-1)
基于 Spark 的标准化 PCA 算法	董建华, 王国胤, 雍 熙, 史晓雨, 李庆亮 (5-7)
基于局部特征的驾驶证自动识别系统	马 岭, 蒋慧琴, 刘玉敏 (5-13)
非接触式红外测温灶具防干烧系统设计	李毅丰, 毛晓波, 杨怡航, 朱 枫 (5-18)
基于发育网络的人脸朝向识别研究	王东署, 谭达佩, 韦晓琴 (5-23)
多传感器信息融合在焊接质量控制中的应用	刘广瑞, 周文博, 田 欣, 郭珂甫 (5-28)

基于谱聚类算法的城市充换电站分布决策	张忠会, 刘故帅, 熊剑峰, 刘小宛, 徐高超 (5-32)
基于粒子群算法的烧结炉系统辨识及神经网络控制	曹 奔, 袁忠于, 刘 洪 (5-39)
基于可能性条件偏好网络的交互式遗传算法及其应用	孙晓燕, 朱利霞, 陈 杨 (6-1)
无线 Ad Hoc 网络中双向传输容量的研究	陈昌海, 李娜娜, 王 静, 李振锋 (6-6)
可重构网络中的一种新型端到端服务供应模型	马 丁, 庄 雷, 兰巨龙, 慕小武 (6-11)
单云服务提供者环境下的随机化属性保护研究	李拴保 (6-17)
基于喷泉码及 CDMA 扩频的数字水印盲算法	张卫党, 吴 琳 (6-23)
基于混合差分演化的网络入侵检测算法	王耀光, 陈伟权, 吴镇邦, 秦 勇, 黄 翰 (6-29)
精英反向学习的单纯形交叉布谷鸟搜索算法	林要华, 王李进 (6-33)

教育管理

高校智慧校园成熟度模型与评价指标体系研究	蒋东兴, 吴海燕, 袁 芳, 付小龙 (2-1)
基于微信公众平台的校园移动信息服务建设方案研究	
.....	龙新征, 欧阳荣彬, 李若森, 李庭晏, 王倩宜 (2-5)
研究生管理信息系统应用效果评价研究	焦宝臣, 陈诗明, 刘振昌, 杨 阳, 张四海 (2-9)
基于属性规则的 PRBAC 参数模型研究与实现	欧阳荣彬, 刘云峰, 龙新征 (2-13)

环境工程

徐州北郊煤矿区优势植物调查及重金属富集特征研究	
.....	张浩嘉, 刘汉湖, 肖 昕, 孙晓菲, 蔡成哲 (2-92)
土壤电参数频变性对接地体有效长度的影响分析	
.....	李景丽, 贺鹏威, 邱再森, 李渊博, 郭丽莹 (4-29)
地温空调井群运行对地下水影响的实验与模型研究	
.....	窦 明, 曹亚新, 米庆彬, 朱志刚, 付 博 (5-76)

本刊被下列数据库及文摘收录:

- 美国《化学文摘》(CA)
- 俄罗斯《文摘杂志》(AJ)
- 英国《科学文摘》(SA)
- 美国《数学评论》(MR)
- 美国《剑桥科学文摘》(CSA)
- 波兰《哥白尼索引》(IC)
- 《中国学术期刊(光盘版)》
- 《中国科技论文统计源》
- 《中文科技期刊数据库》
- 《中国期刊网》
- 《中国学术期刊综合评价数据库》

- 《万方数据系统》
- 《中国机械工程文摘》
- 《中国数学文摘》
- 《中国化学化工文摘》
- 《中国无机分析化学文摘》
- 《中国力学文摘》
- 《中国物理文摘》
- 《工程机械文摘》
- 《中国学术期刊文摘》
- 《电子科技文摘》

郑州大学学报(工学版)
Zhengzhou Daxue Xuebao (Gongxueban)
双月刊 (1980年创刊)
第38卷 第6期 总第156期 2017年11月10日

Journal of Zhengzhou University
(Engineering Science)
Bimonthly (Started in 1980)
Vol.38 No.6 Sum No.156 Nov.10, 2017

主管单位 河南省教育厅
主办单位 郑州大学
出版单位 《郑州大学学报(工学版)》编辑部
电子信箱 gxb@zzu.edu.cn
地 址 河南省郑州市科学大道100号(450001)
电 话 (0371)67781276 67781277
主 编 李燕燕
副 主 编 梁 静
印刷单位 河南金河印务有限公司
发行范围 国内外公开发行
征 订 处 全国各地邮政局
发行单位 河南省邮政发行局

Administrated by The Education Department
of Henan Province
Sponsored by Zhengzhou University
Edited & Published by Editorial Board of Journal of
Zhengzhou University
E-mail gxb@zzu.edu.cn
Address No.100 Science Avenue, Zhengzhou 450001, China
Telephone (0371)67781276 67781277
Chief Editor LI Yanyan
Executive Chief Editor LIANG Jing
Printed by Henan Jinhe Printing Co., Ltd.
Distributed by Publishing Trading Corporation,
P.O.B.782, Beijing 100011, China

中国标准刊号: **ISSN 1671-6833**
CN 41-1339/T

定价: 10.00元

邮发代号: 36-232