基于近似边界和层次聚类的超多目标进化算法

张 峰,顾一凡
(南京航空航天大学 计算机科学与技术学院,江苏 南京 211100)

摘 要:很多工程优化问题需要同时优化超过3个冲突的目标,这类问题就属于超多目标优化问题。由于超多目标优化问题的目标空间过于庞大,并且很多算法往往只能使用数量较少的种群来近似问题的结果,这使得很多算法难以保持较好的多样性和收敛性,此外,许多算法往往忽略使用极值点的有效信息来加速算法收敛。为了解决上述问题,提出了一种基于近似边界和层次聚类的超多目标进化算法。在一种求角点解方法的基础上,使用角点解近似边界(极值点)来加速算法收敛,并进一步提出使用层次聚类来挑选下一代种群,借此使得算法能够保持较好的收敛性和多样性。最后通过与多个流行的求解超多目标优化问题算法进行对比实验,证明了该算法的有效性。

关键词:超多目标优化问题;极值点;超多目标进化算法;角点解;层次聚类

中图分类号:TP301.6 **文献标识码**:A **文章编号**:1673-629X(2020)12-0061-05

doi:10.3969/j.issn.1673-629X.2020.12.011

Many-objective Evolutionary Algorithm Based on Approximate Boundary and Hierarchical Clustering

ZHANG Feng,GU Yi-fan
(School of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics,Nanjing, 211100,China)

Abstract:Many engineering optimization problems need to optimize more than 3 conflicting objectives at the same time,and this type of problem belongs to many-objective optimization problem. The objective space of many-objective optimization problem is too large,and many algorithms can only use a small number of population to approximate the results of the problem,which makes it difficult for many algorithms to maintain better diversity and convergence. In addition,many algorithms often ignore valid information from nadir point to speed up the algorithm's convergence. To solve the above problem,we propose a many-objective evolutionary algorithm based on approximate boundary and hierarchical clustering. On the basis of a corner solution method,the corner solutions is used to approximate the boundary (nadir point) to accelerate the convergence of the algorithm. We further propose the use of hierarchical clustering to select the next population,thereby enabling the algorithm to maintain better convergence and diversity. Finally,the effectiveness of the proposed algorithm is proved by comparing with many popular algorithms for solving super-multi-objective optimization problems.

Key words:many-objective optimization problem;nadir point;many-objective evolutionary algorithm;corner solution;hierarchical clustering

0 引 言

在实际工程优化问题中,往往需要同时优化多个冲突的目标,这类问题就被称为多目标优化问题[1-2](multiobjective optimization problem,MOP)。目标数大于3的多目标优化问题,就被称为超多目标优化问题(many-objective optimization problem,MaOP)。许多工程优化问题都属于超多目标优化问题,高效地求解超多目标优化问题有着十分重要的理论意义和实际价值。

1 概 述

1.1 超多目标优化问题定义

一个连续的最小化超多目标优化问题(MaOP)的数学表达式如下:

$$\begin{aligned} &\text{minimize } F(x) = (f_1(x), \cdots, f_m(x)) \\ &\text{subject to } x \in \Omega \end{aligned} \tag{1}$$

其中,$\Omega \in R^n$ 和 $F:\Omega \rightarrow R^m$ 分别表示决策空间和 m 个实值目标函数组成的目标空间,$x = (x_1, x_2, \cdots, x_n)^T \in \Omega$ 称为MaOP中的一个解。

收稿日期:2020-01-12 **修回日期**:2020-05-15

基金项目:国家自然科学基金(61300159);江苏省自然科学基金(BK20181288)

作者简介:张 峰(1993-),男,硕士研究生,研究方向为智能计算与机器学习。

1.2 相关定义简介

假设两个解 $u,v \in \Omega$,定义 u 支配 v 并简记为 $u \prec v$,成立的条件是当且仅当对 $\forall j \in \{1,2,\cdots,m\}$ 都有 $f_j(u) \leqslant f_j(v)$,并且至少 $\exists i \in \{1,2,\cdots,m\}$ 使得 $f_i(u) < f_i(v)$。设解 $x^* \in \Omega$,若没有解 $x \in \Omega$ 使得 $x \prec x^*$,则称 x^* 和 $F(x^*)$ 分别为 Pareto 最优解和 Pareto 最优向量。所有 Pareto 最优解组成了 Pareto Set(PS),并且 PS 对应的目标向量集合则是 Pareto Front (PF)[3]。

理想点 z^* 的定义如下:

$$z_j^* = \min_{x \in \Omega} f_j(x),\ j \in \{1,2,\cdots,m\} \tag{2}$$

极值点 z^{nad} 的定义如下:

$$z_j^{nad} = \max_{x \in P'} f_j(x),\ j \in \{1,2,\cdots,m\} \tag{3}$$

其中,P' 表示一个解集 P 的非支配解集。

决策者可以根据实际需要从算法近似的 PF 中挑选一个解进行决策。

收敛性和多样性是衡量多目标优化算法性能的准则。收敛性是指算法近似出来的 PF 距离真实 PF 的远近,多样性是指算法近似的 PF 能否很好地覆盖和表示整个 PF。

1.3 多目标优化算法研究综述

在过去的几十年中,许多经典的多目标优化算法被提出来求解 MOP,而多目标进化算法能够通过进行大量目标函数评估来有效求解 MOP。下面简单介绍一些多目标进化算法。

NSGA-II[4] 能够通过快速非支配排序和计算拥挤距离来有效求解 2 到 3 目标的 MOP,但在求解 MaOP 上,会出现许多非支配解,导致算法性能难以达到理想的效果。MOEA/D-DE[5] 通过结合基于一组权重向量分解的方法和差分进化来有效求解一些具有复杂 PS 形状的 MOP,但由于算法中的权重向量是固定的,这不一定能很好地适应 PF 形状,在求解一些 PF 形状特殊的 MaOP,MOEA/D-DE 效果欠佳。为了更好地平衡算法在高维目标空间上的收敛性和多样性,RVEA[6] 通过一种角度惩罚距离的度量方法来自适应地调整权重向量分布,因此,RVEA 能够较好地求解 MaOP 问题。在一种被广泛使用的评价指标 Inverted Generational Distance[7](IGD)上,MaOEA/IGD[8] 设计了一种高效计算支配比较的方法和提出了一种基于分解的高效近似极值点的方法,来更好地求解 MaOP。尽管上述的多目标进化算法能够从基于支配、分解和指标的角度来较好地求解 MaOP 问题,但由于 MaOP 目标空间过于庞大,算法使用的种群数量有限,如何在求解 MaOP 时更好地考虑算法的收敛性和多样性一直是一个严峻的挑战。

1.4 使用角点解近似理想点和极值点

Singh 和 Isaacs 等人定义了角点解并提出一种求出角点解的方法[9]。角点解可能出现在以下这两种极端的情况。(1)角点解位于一个只有某个目标最小的超平面上;(2)角点解位于轴坐标上,原点在理想点上。

角点解可以用来有效近似极值点。假设 S_C 是一个角点解集,极值点的近似公式如下所示:

$$z_j^{nad} \approx \max_{x \in S_C} f_j(x),\ j \in \{1,2,\cdots,m\} \tag{4}$$

1.5 研究动机

在求解 MOP 时,平衡多样性和收敛性是提升算法性能的关键。但在求解 MaOP 时,由于目标空间过大,并且算法往往只能使用数量较少的种群来近似问题的 PF,这使得许多算法更难以较好地保持多样性和收敛性。此外,大多数算法忽略了使用近似的极值点来固定近似的 PF 边界,这会影响 MaOP 的求解速度。为了充分利用极值点的有效信息,该文使用一种求角点解的方法[9]来近似极值点,和使用极值点固定近似的 PF 边界来排除比较差的解,借此来加速算法收敛,然后进一步提出使用层序聚类的方法来将种群划分成多个聚类,并根据每个聚类中心自适应地产生一个权重向量,最后根据每个聚类中的解在对应权重向量的聚合函数来挑选下一代种群。通过上述方法使得算法能够保持较好的收敛性和多样性并有效地求解 MaOP。

2 算法设计

2.1 算法框架

本节主要介绍提出的基于层次聚类和边界近似的超多目标进化算法(MaOEA-ABHC)的实现细节,其中 MaOEA 表示超多目标进化算法,ABHC 表示边界近似和层次聚类。MaOEA-ABHC 的算法框架如算法 1 所示:

算法 1:MaOEA-ABHC 算法框架。

输入:N:种群数;

算法的终止条件。

输出:最终种群 P。

1. 随机生成规模为 N 的初始种群 P;
2. while 没达到终止条件 do
3. 对 P 使用 SBX 算子产生子代 Q;
4. $P_U = P \cup Q$;
5. 使用 P_U 近似 z^*;
6. 根据 P_U 求出角点解集 S_C 并近似 z^{nad};
7. P = UpdatePop(N, P_U, S_C, z^*, z^{nad});
8. end
9. return P

MaOEA-ABHC 主要分为 3 个阶段:

(1)初始化:主要随机生成一个规模为 N 的初始种群 P 。

(2)近似边界:首先对当前种群 P 使用 SBX 算子来产生子代 Q ,然后根据式(5)来近似理想点 z^* :

$$z_j^* \approx \min_{x \in P_U} f_j(x),\ j \in \{1,2,\cdots,m\} \tag{5}$$

根据 Singh 和 Isaacs 等人的工作[9]求出当前 P_U 的角点解集 S_C ,最后根据式(4)使用 S_C 来近似极值点 z^{nad} 。

(3)更新种群:根据 P_U 非支配解的数量与种群数的相应情况来更新当前种群 P 。

2.2　更新种群

更新种群相对比较复杂,首先获取 P_U 的非支配解集 P' 和剩余解集 P_R 。

接下来进行第一次判断,如果 P' 的数量 $|P'|$ 大于种群数 N ,则先去除 P' 在 z^{nad} 外的解,并得到剩余解集 P' 和去除解集 P_E ,再进行第二次判断,如果 P' 的数量 $|P'|$ 仍然大于种群数 N ,则使用层次聚类方法来挑选下一代种群。第二次判断中,如果 P' 的数量 $|P'|$ 小于种群数 N ,则根据 P_E 中的解到 z^* 的最短欧氏距离来挑选 $N - |P'|$ 个解补充到 P' 。第一次判断中,如果 P' 的数量 $|P'|$ 小于种群数 N ,则根据 P_R 中的解到 z^* 的最短欧氏距离来挑选 $N - |P'|$ 个解补充到 P' 。最后返回一个新种群 P' ,具体的实现细节可以参考算法 2。

算法 2:UpdatePop。

输入: N :种群数; P_U :父代和子代的合并种群; S_C : P_U 的角点解集; z^* :近似的理想点; z^{nad} :近似的极值点。

输出:新种群 P' 。

1. 获取 P_U 的非支配解集 P' 和剩余解集 P_R ;
2. if $|P'| > N$ then
3. 去除 P' 在 z^{nad} 外的解,得剩余解集 P' 和去除解集 P_E ;
4. if $|P'| > N$ then
5. P' =HCSelection(N , P' , S_C);
6. else
7. 根据 P_E 中的解到 z^* 的欧氏距离挑选 $N - |P'|$ 个解补充到 P' ;
8. end
9. else
10. 根据 P_R 中的解到 z^* 的欧氏距离挑 $N - |P'|$ 个解补充到 P' ;
11. end
12. return P'

2.3　层次聚类选解

首先把 P_U 的角点解集 S_C 加入到新种群 P ,接下来使用层次聚类的方法挑选出剩下的 $N - |S_C|$ 个解加入到 P 中, $|S_C|$ 表示角点解集的数量。

在层词聚类前先对每个解 $x \in P'$ 的目标向量进行归一化,归一化的公式如下:

$$\frac{f_j(x) - f_j^{\min}}{f_j^{\max} - f_j^{\min}},\ j \in 1,2,\cdots,m \tag{6}$$

其中, $f_j^{\max}$ 和 $f_j^{\min}$ 分别表示 P' 中的解在第 j 个目标函数的最大值和最小值。

接下来,先对 P' 中每个解的目标向量进行归一化后得到 P'_1 ,并把 P'_1 的目标向量变成单位向量后得到 P'_2 。根据 P'_2 的目标向量使用层次聚类把 P'_2 划分为 $N - |S_C|$ 个聚类,每个聚类的解集索引号保存到 I 。在 I 和 P'_1 的基础上,计算每个聚类 $P'_1(\mathrm{idx})$ ($\mathrm{idx} \in I$)的聚类中心 c^{idx} ,并根据这个聚类中心自适应地产生一个权重向量 λ^{idx} 。第 $i \in 1,2,\cdots,N-|S_C|$ 个聚类中心 $c^i = (x_1^i, x_2^i, \cdots, x_n^i)^{\mathrm{T}}$ 产生权重向量 $\lambda^i = (\lambda_1^i, \lambda_2^i, \cdots, \lambda_n^i)^{\mathrm{T}}$ 的具体公式如下:

$$\lambda_j^i = \frac{x_j^i}{x_1^i + x_2^i + \cdots + x_n^i},\ j \in 1,2,\cdots,n \tag{7}$$

最后根据 MOEA/D[10] 提出的 Weighted Sum (WS)方法计算 $P'_1(\mathrm{idx})$ 中的解在 λ^{idx} 下的聚合函数,并从 $P'_1(\mathrm{idx})$ 中挑选一个最好的解 x^B 进行评估,然后把 x^B 加入到 P 中。HCSelcection 函数最后会返回一个新种群 P ,具体的细节可以参考算法 3。

算法 3:HCselection。

输入: N :种群数; P' :候选种群; S_C : P_U 的角点解集。

输出:新种群 P 。

1. $P = \varnothing$;
2. $P = P \cup S_C$;
3. 按式(6)对 P' 中的目标向量进行归一化,得 P'_1 ;
4. 把 P'_1 中的目标向量变成单位向量,得 P'_2 ;
5. 对 P'_2 进行层次聚类,得一组聚类索引 I ;
6. foreach idx $\in I$ do
7. 计算解集 $P'_1(\mathrm{idx})$ 的聚类中心 c^{idx} ;
8. 按式(7)来计算聚类中心权重向量 λ^{idx} ;
9. 按 WS 计算 $P'_1(\mathrm{idx})$ 在 λ^{idx} 的聚合函数;
10. 获取并评估 $P'_1(\mathrm{idx})$ 中的最佳个体 x^B ;
11. $P = P \cup x^B$;
12. end
13. return P

3　实验设计

3.1　实验参数设置

MaF 系列测试问题[11]是一组流行的超多目标优化问题。该文主要选择了 MaF 系列中前 11 个测试问题来检验 MaOEA-ABHC 与 4 个流行超多目标优化对比算法的性能,这些对比算法主要包括 NSGA-II、MOEA/D-DE、RVEA 和 MaOEA/IGD。每个算法都

独立地运行目标变量数 m 为 5 和 10 的 MaF1-MaF11 测试问题 30 次。在 5 目标的 MaF 中,除了 MaF7 的决策变量数 n 为 24 和 MaF8-9 的 n 为 2 外,其他问题的 n 都为 14;在 10 目标的 MaF 中,除了 MaF7 的决策变量数 n 为 29 和 MaF8-9 的 n 为 2 外,其他问题的 n 都为 19。所有算法的种群数都设为 200,并且每个测试问题的最大目标函数评估次数都设为 100 000 次。

除了上文提到的 IGD 指标外,还有 Hypervolume Indicator[12](HV)、IGD+[13]和 R_2 [14]等评价指标。该文主要使用一种考虑收敛性和多样性的综合指标 HV 来评价所有算法的性能,以上算法都在 PlatEMO[15]上实现,并使用 PlatEMO 自带的 HV 计算方法来计算所有算法的 HV 值。此外,还使用了置信度为 0.05 的 Wilcoxon 秩和检验来度量 MaOEA-ABHC 与其他对比算法的显著性,其中符号=表示 MaOEA-ABHC 和其他对比算法没有显著性区别,符号+和-分别表示 MaOEA-ABHC 明显地比其他对比算法好和差。

3.2 实验结果

正如表 1 所示,MaOEA-ABHC 在大多数测试问题上都取得了最佳的 HV 值,并且在大多数测试问题上,MaOEA-ABHC 的算法性能都显著地优于其他对比算法。由于 HV 是一种考虑多样性和收敛性的综合指标,这表明 MaOEA-ABHC 在解决各类型的 MaOP 上都能保持较好的多样性和收敛性,并能有效地求解各类型的 MaOP。

表 1 所有算法在全部测试问题上的 HV 均值和显著性符号(每个问题的最佳指标值加粗显示)

测试问题	M	MaOEA-ABHC	NSGA-II	MOEA/D-DE	MaOEA-IGD	RVEA
MaF1	5	**3.238 6e-2**	2.668 8e-2 +	2.187 8e-2 +	7.984 3e-3 +	1.088 8e-2 +
MaF2	5	**2.645 5e-1**	2.367 7e-1 +	2.053 3e-1 +	2.050 6e-1 +	2.403 4e-1 +
MaF3	5	**9.986 5e-1**	5.077 8e-1 +	8.086 3e-1 +	2.097 2e-1 +	9.980 8e-1 =
MaF4	5	**1.829 4e-1**	1.452 7e-1 +	3.893 2e-2 +	1.062 4e-2 +	4.905 5e-2 +
MaF5	5	**8.678 5e-1**	7.953 7e-1 +	6.822 6e-1 +	6.942 3e-1 +	8.466 8e-1 +
MaF6	5	2.104 9e-1	**2.105 5e-1 =**	1.969 2e-1 +	1.343 3e-1 +	1.917 3e-1 +
MaF7	5	**3.428 7e-1**	2.957 7e-1 +	1.177 3e-1 +	2.602 1e-1 +	2.958 5e-1 +
MaF8	5	**1.746 8e-1**	1.639 7e-1 +	1.632 7e-1 +	8.081 6e-2 +	1.108 4e-1 +
MaF9	5	**3.786 7e-1**	2.376 2e-1 +	3.826 3e-1 =	1.533 5e-1 +	2.841 6e-1 +
MaF10	5	9.683 5e-1	9.644 8e-1 =	9.480 6e-1 +	4.560 8e-1 +	**9.853 5e-1 -**
MaF11	5	**9.937 5e-1**	9.915 1e-1 +	9.434 5e-1 +	9.456 4e-1 +	9.933 9e-1 =
MaF1	10	**1.876 3e-5**	6.900 0e-6 +	1.248 9e-6 +	2.072 0e-6 +	5.813 6e-7 +
MaF2	10	**2.852 4e-1**	2.577 8e-1 +	2.434 0e-1 +	2.169 3e-1 +	1.344 1e-1 +
MaF3	10	**9.996 4e-1**	0.000 0e+0 +	9.981 5e-1 +	4.212 6e-1 +	9.844 8e-1 +
MaF4	10	**1.131 6e-3**	1.711 5e-4 +	8.831 8e-6 +	7.269 2e-7 +	4.742 1e-6 +
MaF5	10	**9.762 6e-1**	1.652 1e-3 +	5.255 7e-1 +	5.586 5e-1 +	9.021 3e-1 +
MaF6	10	**1.777 0e-1**	1.310 1e-2 +	1.748 1e-1 +	1.002 5e-1 +	1.687 1e-1 +
MaF7	10	**2.414 5e-1**	4.225 5e-5 +	1.961 9e-4 +	4.895 6e-2 +	2.323 8e-1 +
MaF8	10	**2.012 6e-2**	1.676 3e-2 +	1.564 3e-2 +	4.304 7e-3 +	5.701 1e-3 +
MaF9	10	**3.132 7e-2**	1.088 1e-4 +	2.260 4e-2 +	7.784 5e-3 +	7.908 4e-3 +
MaF10	10	**9.835 9e-1**	9.202 9e-1 +	4.602 0e-1 +	7.278 2e-1 +	9.803 0e-1 +
MaF11	10	**9.984 7e-1**	9.973 2e-1 +	9.898 6e-1 +	9.264 4e-1 +	9.641 3e-1 +

4 结束语

许多工程优化问题都属于超多目标优化问题。由于目标空间过于庞大,并且算法往往只能使用规模较少的种群来近似问题的 PF,这使得多数算法难以保持较好的收敛性和多样性。此外,许多算法往往忽视使用极值点的有用信息来帮助算法加速收敛并提升优化性能。在一个求角点解集方法的基础上,该文使用角点解集近似边界(极值点)并加速算法收敛,然后使用层次聚类进行选解来保证算法的收敛性和多样性。在未来,将进一步拓展算法来研究并求解一些实际应用问题。

参考文献:

[1] DEB K. Multiobjective Optimization Using Evolutionary Algorithms[M]. Chichester,U. K.:Wiley,2001.

[2] 康　锰,许　峰. 多进化策略自适应免疫多目标进化算法[J]. 安徽理工大学学报:自然科学版,2019,39(5):43-47.

[3] MIETTINEN K. Nonlinear Multiobjective Optimization[M]. Norwell,MA:Kluwer,1999.

[4] DEB K,PRATAP A,AGARWAL S,et al. A Fast And Elitist Multiobjective Genetic Algorithm: NSGA - II [J]. IEEE Transactions On Evolutionary Computation,2002,6(2):182-197.

[5] LI H, ZHANG Q. Multiobjective Optimization Problems With Complicated Pareto Sets, MOEA/D And NSGA - II [J]. IEEE Transactions On Evolutionary Computation,2008,13(2):284-302.

[6] CHENG R,JIN Y,OLHOFER M,et al. A Reference Vector Guided Evolutionary Algorithm For Many-Objective Optimization[J]. IEEE Transactions On Evolutionary Computation,2016,20(5):773-791.

[7] BOSMAN P A N, THIERENS D. The Balance Between Proximity And Diversity In Multiobjective Evolutionary Algorithms[J]. IEEE Transactions on Evolutionary Computation,2003,7(2):174-188.

[8] SUN Y,YEN G G,YI Z. IGD indicator-based evolutionary algorithm for many - objective optimization problems[J]. IEEE Transactions on Evolutionary Computation, 2018, 23(2):173-187.

[9] SINGH H K,ISAACS A,RAY T. A Pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems[J]. IEEE Transactions on Evolutionary Computation,2011,15(4):539-556.

[10] ZHANG Q,LI H. MOEA/D: a multiobjective evolutionary algorithm based on decomposition[J]. IEEE Transactions on Evolutionary Computation,2007,11(6):712-731.

[11] CHENG R,LI M,TIAN Y,et al. A benchmark test suite for evolutionary many - objective optimization[J]. Complex & Intelligent Systems,2017,3(1):67-81.

[12] ZITZLER E,THIELE L. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach[J]. IEEE Transactions on Evolutionary Computation,1999,3(4):257-271.

[13] ISHIBUCHI H,MASUDA H,TANIGAKI Y,et al. Modified distance calculation in generational distance and inverted generational distance[C]//International conference on evolutionary multi-criterion optimization. Cham:Springer,2015:110-125.

[14] BROCKHOFF D,WAGNER T,TRAUTMANN H. On the properties of the R2 indicator[C]//Proceedings of the 14th annual conference on genetic and evolutionary computation. Philadelphia,PA:ACM,2012:465-472.

[15] TIAN Y,CHENG R,ZHANG X,et al. PlatEMO: a MATLAB platform for evolutionary multi-objective optimization [educational forum][J]. IEEE Computational Intelligence Magazine,2017,12(4):73-87.

(上接第 60 页)

[2] FRANKLIN M,HALEVY A,MAIER D. From databases to dataspaces[J]. ACM SIGMOD Record,2005,34(4):27-33.

[3] BLUNSCHI L,DITTRICH J P,GIRARD O R,et al. A dataspace Odyssey: the iMeMex personal dataspace management system[C]//Proceedings of the 3rd conference on innovative data systems research (CIDR 2007). Asilomar,USA:[s. n.],2007:114-119.

[4] DONG Xin,HALEVY A. A platform for personal information management and integration[C]//Proceedings of the 2nd conference on innovative data system research (CIDR 2007). New York:ACM,2005:119-130.

[5] LI Y,MENG X. Research on personal dataspace management[C]//International conference on management of data. Vancouver,Canada:ACM,2008.

[6] DITTRICH J P,SALLES M A V. iDM: a unified and versatile data model for personal dataspace management[C]//International conference on very large data bases(VLDB'06). Seoul,Korea:[s. n.],2006:367-378.

[7] ZHONG Ming,LIU Mengchi,CHEN Qian. Modeling heterogeneous data in dataspace[C]//IEEE international conference on information reuse & integration (IEEE IRI 2008). Las Vegas,Nevada,USA:IEEE,2008:404-409.

[8] 董彦磊,申德荣,寇　月,等. 数据空间中数据组织模型以及关联关系发现模型的研究[J]. 计算机研究与发展,2009,46(Suppl.):191-199.

[9] 杨　丹,申德荣,聂铁铮,等. 数据空间中数据模型及实体关联关系挖掘的研究[J]. 小型微型计算机系统,2012,33(5):936-939.

[10] 王江海,武林仙,吴扬扬. 基于刻面的数据空间数据源管理子系统[J]. 华侨大学学报:自然科学版,2012,33(5):509-512.

[11] 李玉坤,任　标,赵喜燕,等. 个人数据管理技术研究[J]. 计算机科学与探索,2014,8(11):1281-1295.

[12] 刘正涛,王建东. Web 数据空间技术研究[J]. 计算机工程与应用,2012,48(7):12-19.

[13] 周文涛. 一种企业数据空间可视化汇聚流程建模方法与查询优化策略[D]. 青岛:山东科技大学,2010.

[14] VAN BRUGGEN R,BATON J. Learning Neo4j 3. x - second edition[M]. Birmingham:Packt Publishing,2017:33-36.

[15] 贲福才. 统一大块数据存取方法的研究[D]. 大庆:大庆石油学院,2010.

[16] 马殿富,郎　波,黄　雷,等. 非结构化数据表示规范[S]. 北京:中国国家标准化管理委员会,2016.