第 32 卷 第 8 期
2009 年 8 月

计 算 机 学 报
CHINESE JOURNAL OF COMPUTERS

Vol. 32 No. 8
Aug. 2009

GPP 问题的骨架分析与启发式算法设计

江 贺[1),2)] 邱 铁[1)]

[1)](大连理工大学软件学院 大连 辽宁 116621)

[2)](中国科学院软件研究所计算机科学国家重点实验室 北京 100190)

摘 要 图的划分问题(GPP)是具有广泛应用背景的典型 NP-难解问题,高效启发式算法一直是该领域的研究热点.作为设计启发式算法的有力工具,GPP 的骨架分析存在理论分析结果匮乏、骨架规模过小等缺陷.文中采用构造偏移 GPP 实例的技巧,不仅在理论上证明了获取 GPP 的骨架是 NP-难解的,并且利用一般 GPP 实例与偏移实例的关系,实现了骨架规模的提高.在此基础上,文中对于目前求解 GPP 问题最好的算法之一的 IBS 进行了改进,提出了基于偏移实例的 IBS 算法(BI-IBS).算法 BI-IBS 首先构造偏移 GPP 实例,然后再利用局部最优解交集对它进行归约,最后再求解归约后的规模更小的新实例.实验结果表明,BI-IBS 比现有算法在解的质量上有了较显著的提高.文中的工作较完善地解决了 GPP 的骨架研究存在的问题,所采用的构造偏移实例的技巧对于其它 NP-难解问题的骨架理论分析及启发式算法设计亦具有较高的参考价值.

关键词 图的划分问题;NP-难解;骨架分析;启发式算法设计

中图法分类号 TP18 **DOI 号**: 10.3724/SP.J.1016.2009.01662

Backbone Analysis and Heuristic Design for the Graph Partitioning Problem

JIANG He[1),2)] QIU Tie[1)]

[1)](*School of Software, Dalian University of Technology, Dalian, Liaoning* 116621)

[2)](*State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing* 100190)

Abstract The graph partitioning problem (GPP) is one of the typical NP-Hard problems with extensively wide applications. Efficient heuristics have been the hotline in this research area at all times. There exist defects of theoretic results and limited size in backbone analysis, a useful tool for heuristic design of the GPP. In this paper, by constructing the biased instance, it is proved that it is NP-Hard to obtain the GPP backbone, and the backbone size is increased through analyzing the relationship between the biased GPP instance and general GPP instances. Furthermore, the biased instance-IBS (BI-IBS) is proposed to improve the IBS which is one of the best heuristics for the GPP. In this new heuristic, after a biased GPP instance was built and reduced by intersection of local optimal solutions, the reduced instance was then solved. Experiments indicate that the BI-IBS had achieved better performance than existing heuristics in terms of solution quality. Not only this work solved existing problems in research of the GPP backbone, but the skill of biased instance construction threw a light on the theoretic analysis of backbone and heuristic design for other NP-Hard problems.

Keywords graph partitioning problem; NP-hard; backbone analysis; heuristic design

---

收稿日期:2007-04-02;最终修改稿收到日期:2009-05-31.本课题得到国家自然科学基金(60805024)、教育部博士点基金(20070141020)资助.**江 贺**,男,1980 年生,博士,副教授,研究方向为智能计算、数据挖掘等.E-mail: jianghe@dlut.edu.cn. **邱 铁**,男,1980 年生,博士研究生,研究方向为数据挖掘.

1 引　言

图的划分问题 GPP（Graph Partitioning Problem）[1]属于典型的 NP-难解问题，在并行计算、VLSI 设计、交通调度和数据挖掘等众多领域具有极其广泛的应用背景. 根据计算复杂性理论：除非 P＝NP，否则 NP-难解问题不存在多项式时间的完全算法. 故目前对于 GPP 的算法研究包括近似算法（Approximate Algorithm）[1-2]和启发式算法（Heuristic）. 对于大规模实例，人们的研究目标集中在能用较短时间得到可接受（在性能上）解的启发式算法方面. 现有的求解 GPP 问题的启发式算法分为简单启发式算法和元启发式算法两类：前者包括 Kernighan-Lin 算法[3]、贪心算法[4]；后者包括模拟退火算法[5]、禁忌算法[6]、遗传算法[7]、Memetic 算法[8]、Multilevel 算法[9].

骨架（backbone）作为提高启发式算法性能的有效工具，是当前算法研究的前沿领域[10-12]. 骨架是指一个 NP-难解问题实例的所有全局最优解的相同部分. 人们利用骨架或者部分骨架对算法的搜索空间进行收缩，从而显著提高了算法性能. Schneider[13]等利用局部最优解交集作为近似骨架，得到了求解旅行商问题 TSP（Traveling Salesman Problem）的多级算法；Zhang[14]、Dubois[15]等分别给出求解约束可满足性问题 SAT（the Satisfiability Problem）的骨架导向局部搜索；Zou 等[16]提出了求解二次分配问题 QAP（Quadratic Assignment Problem）的近似骨架导向的蚁群算法.

目前，GPP 问题的骨架研究还处于起步阶段，仅有的结果是：Zou 等[17]利用局部最优解交集来模拟骨架，由此给出的 IBS 算法（Intersection-Based Sizescaling）在求解质量方面较已有的主流算法 Kernighan-Lin[3]和 Multilevel 算法[9]有较显著的提高. 然而，目前 GPP 问题的骨架研究还存在两大缺陷：(1) 骨架分析还采用实验统计的手段，缺乏理论分析结果. 在我们的知识范围内，NP-难解问题的骨架研究仅有的理论分析是在 TSP 问题上，由 Kilby 等[18]证明了获取 TSP 的骨架属于 NP-难解的. 而骨架的理论分析对于启发式算法设计极其重要：若可以在多项式时间内得到骨架，则可以直接使用骨架而非近似骨架. (2) IBS 算法无法解决骨架规模过小的难题. IBS 使用局部最优解交集来逼近骨架，而当全局最优解个数不唯一时，骨架的规模可能非常小，导致近似骨架难以发挥其应有的作用.

针对上述缺陷，本文通过构造偏移实例的技巧，同时解决了骨架理论分析与骨架规模提高的难题. 论文首先给出了偏移实例的定义，并证明了其具有两大优点：(1) 恰好只有唯一的全局最优解；(2) 从骨架可以构造出全局最优解. 随后，论文分析了一般 GPP 实例与偏移实例的关系：任意 GPP 实例均可以构造出相应的偏移实例，而偏移实例的全局最优解恰好是原实例的一个全局最优解，这样就实现了骨架规模的提高，进而从理论上证明了获取 GPP 的骨架是 NP-难解的. 因此，在 P≠NP 的前提下，寻求获取骨架的多项式时间算法是没有意义的. 在此基础上，论文利用偏移实例的构造方法，给出了基于偏移实例的 IBS 算法 BI-IBS（Biased Instance-IBS）. BI-IBS 算法首先构造偏移 GPP 实例，然后再利用局部最优解交集对它进行归约，最后再求解归约后的规模更小的新实例. 论文以 Kernighan-Lin 算法为从属算法，对 BI-IBS、IBS、Multilevel 进行了测试，实验结果表明，BI-IBS 算法在解的质量上较 IBS、Multilevel 有较为明显的提高.

本文的工作不仅拓广了 GPP 问题的研究内容，骨架理论分析中所采用的构造偏移实例的方法对于其它 NP-难解问题亦具有较高的参考价值. 而论文中提出的先构造偏移实例，再利用其获得近似骨架的方法，为其它 NP-难解问题的启发式算法设计提供了一种新的思路.

2 预备知识

本节首先给出文中所用到的一些定义和记号.

定义 1. 给定无向加权图 $G=(V,E,c,w)$，其中 $V=\{1,2,\cdots,n\}$ 为顶点集合，E 表示无向边的集合（即 (i,j)，(j,i) 表示同一条边），c_i 表示顶点 i 的权值，w_{ij} 表示边 (i,j) 的权值. 记 $|V|=n$ 表示集合 V 中顶点数目，GPP 实例（记为 $GPP(V,E,c,w)$）的可行解 $s=\{V_1,V_2,\cdots,V_k\}$ 将 V 划分为 k 个互不相交的子集 $V_1,V_2,\cdots,V_k$（k 为预先定义的数值，$1\le k\le n$），且满足 $\sum_{i\in V_j} c_i \le \lceil \sum_{i\in V} c_i/k \rceil$ $(1\le j\le k)$①.

定义 2. 给定 GPP 实例 $GPP(V,E,c,w)$，对于可行解 $s=\{V_1,V_2,\cdots,V_k\}$，记 $C(s)=\{(i,j)\mid (i,j)\in E, i\in V_p, j\in V_q, 1\le p<q\le k\}$ 表示解 $s=\{V_1,V_2,\cdots,V_k\}$ 内各个顶点子集间的割（cut），可行解 s 的目标函数值定义为 $w(s)=\sum_{(i,j)\in C(s)} w_{ij}$. GPP

① $\lceil\cdot\rceil$ 表示上取整函数，例如 $\lceil 1.5\rceil=2$，$\lceil 1.0\rceil=1$.

问题的目标是寻求目标函数值最小的可行解 s^*，即 $w(s^*)=\min(\{w(s)|s\in\Pi\})$，其中 Π 是全体可行解的集合.

定义 3. 给定无向加权图 $G=(V,E,c,w)$，定义偏移边集合 $\hat{E}=\{(i,j)|i,j\in V,i\neq j\}$，定义偏移边权值为 $\hat{w}_{ij}=\begin{cases}w_{ij}+1/2^{in+jn+|i-j|}, & (i,j)\in E\\ 1/2^{in+jn+|i-j|}, & (i,j)\notin E\end{cases}$.
本文称 $GPP(V,\hat{E},c,\hat{w})$ 为 $GPP(V,E,c,w)$ 的偏移 GPP 实例. 对于 $GPP(V,\hat{E},c,\hat{w})$ 的可行解 $s=\{V_1,V_2,\cdots,V_k\}$，它的割记为 $\hat{C}(s)=\{(i,j)|(i,j)\in\hat{E},i\in V_p,j\in V_q,1\le p<q\le k\}$，目标函数值 $\hat{w}(s)=\sum_{(i,j)\in\hat{C}(s)}\hat{w}_{ij}$.

容易验证，对于任意的两条边 (i,j)，(m,l)，除非(1) $i=m,j=l$ 或者(2) $i=l,j=m$，否则 $in+jn+|i-j|\neq mn+ln+|m-l|$. 显然，根据定义 3 可知下面性质成立.

性质 1. 给定无向加权图 $G=(V,E,c,w)$，偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$ 的任意可行解也是 $GPP(V,E,c,w)$ 的可行解.

定义 4. 给定 GPP 实例 $GPP(V,E,c,w)$，存在有限多个全局最优解 $s_1^*,s_2^*,\cdots,s_r^*$，记所有全局最优解的集合为 Π^*，其中 $r=|\Pi^*|$ 为全局最优解个数. 实例 $GPP(V,E,c,w)$ 的骨架定义为 $bone(V,E,c,w)=C(s_1^*)\cap C(s_2^*)\cap\cdots\cap C(s_r^*)$.

定义 5. 给 GPP 实例 $GPP(V,E,c,w)$，若其有且仅有一个全局最优解，则称该实例是唯一全局最优解实例.

3 GPP 问题的骨架分析

本节主要分析 GPP 问题的骨架的计算复杂性. 首先，本节将证明偏移 GPP 实例是唯一全局最优解实例，然后再说明偏移 GPP 实例与一般 GPP 实例在解之间的关系，并由此证明获取 GPP 问题的骨架是 NP-难解的.

定理 1. 给定 GPP 实例 $GPP(V,E,c,w)$，若 $w_{ij}\in Z((i,j)\in E)$，则它的偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$ 是唯一全局最优解实例.

证明. 对于 $GPP(V,\hat{E},c,\hat{w})$ 的任意两个不同的解 $s_1=\{V_1^1,V_2^1,\cdots,V_k^1\}$ 和 $s_2=\{V_1^2,V_2^2,\cdots,V_k^2\}$，由 $s_1\neq s_2$ 可知 $\hat{C}(s_1)\neq\hat{C}(s_2)$，故存在 $(i^*,j^*)\in\hat{C}(s_1)$ 且 $(i^*,j^*)\notin\hat{C}(s_2)$. 在使用二进制表示权值的情况下，$\hat{w}(s_1)$ 在小数位的第 $i^*n+j^*n+|i^*-j^*|$ 位为 1，而 $\hat{w}(s_2)$ 该位为 0，故 $\hat{w}(s_1)\neq\hat{w}(s_2)$，原命题得证.

证毕.

下面，本文说明割与可行解之间的关系. 给定无向加权图 $G=(V,E,c,w)$，若图 G 是完全图，给定实例 $GPP(V,E,c,w)$ 的可行解的割，可以如算法 1 构造可行解.

算法 1. 割的可行解构造算法.

输入：GPP 实例 $GPP(V,E,c,w)$，割 E_C，k

输出：解 s

Begin

1. $V'=V,s=\varnothing$；
2. for $t=1$ to k do
   2.1. 取 $i\in V'$；
   2.2. $V_t=\{i\}$；
   2.3. $V'=V'-i$；
   2.4. for each $j\in V'$ do
        if $(i,j)\notin E_C$ then $V'=V'-j$, $V_t=V_t\cup\{j\}$；
   2.5. $s=s\cup\{V_t\}$；

End

算法 1 的正确性是显而易见的：完全图上任意两个顶点间均存在边相连，若顶点 i,j 分属于不同的子集，则必有边 $(i,j)\in E_C$ 成立；反之，若 $(i,j)\notin E_C$，则顶点 i,j 必属于同一个子集. 算法步 1 的时间复杂度为 $O(n)$；步 2.1～2.3 和步 2.5 的时间复杂度均为 $O(1)$，步 2.4 的时间复杂度为 $O(n)$，故步 2 的时间复杂度为 $O(kn)$. 算法 1 的时间复杂度为 $O(kn+n)$.

引理 1. 给定 GPP 实例 $GPP(V,E,c,w)$，若 $w_{ij}\in Z((i,j)\in E)$，给定任意两个不同的可行解 $s_1=\{V_1^1,V_2^1,\cdots,V_k^1\}$ 和 $s_2=\{V_1^2,V_2^2,\cdots,V_k^2\}$，若 $w(s_1)<w(s_2)$，则偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$ 满足 $\hat{w}(s_1)<\hat{w}(s_2)$ 成立.

证明. 对于 GPP 实例 $GBP(V,E,c,w)$ 的任意两个不同的可行解 s_1 和 s_2，若 $w(s_1)<w(s_2)$，根据 $w_{ij}\in Z((i,j)\in E)$ 可知 $w(s_2)-w(s_1)\ge 1$.

而根据偏移 GPP 实例的定义可知：

$$0<\hat{w}(s_1)-w(s_1)=\sum_{(i,j)\in\hat{C}(s_1)}1/2^{in+jn+|i-j|}\le\sum_{(i,j)\in\hat{C}(s_1)}1/2^{3n}\le(n(n-1)/2)(1/2^{3n})<n^2/2^{3n+1}<1.$$

同理可知，$0<\hat{w}(s_2)-w(s_2)<1$. 故有以下等式成立：

$$\hat{w}(s_2)-\hat{w}(s_1)=w(s_2)-w(s_1)+(\hat{w}(s_2)-w(s_2))-(\hat{w}(s_1)-w(s_1))>w(s_2)-w(s_1)-1\ge 0.$$

原命题得证. 证毕.

引理 2. 给定 GPP 实例 $GPP(V,E,c,w)$，若

$w_{ij}\in Z((i,j)\in E)$,则偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$ 的全局最优解也是 $GPP(V,E,c,w)$ 的全局最优解.

证明. 首先,$GPP(V,\hat{E},c,\hat{w})$ 的全局最优解也是 $GPP(V,E,c,w)$ 的一个可行解. 下面用反证法证明,假设实例 $GPP(V,\hat{E},c,\hat{w})$ 的全局最优解 s^* 不是 $GPP(V,E,c,w)$ 的全局最优解,则存在一个解 s,使得 $w(s)<w(s^*)$ 成立. 那么根据引理 1 可知 $\hat{w}(s)<\hat{w}(s^*)$ 成立,矛盾,原命题得证. 证毕.

定理 2. 在 P≠NP 的假设下,不存在多项式时间算法可以获得 GPP 问题的骨架.

证明. 反证法. 假设原命题不成立,则存在某个算法 Λ 可以在多项式时间(记为 $O(\Diamond)$)内得到 GPP 问题的骨架. 下面将构造多项式时间算法获得其全局最优解,从而说明矛盾的存在.

给定任意的 GPP 实例 $GPP(V,E,c,w)$(不妨假定 $w_{ij}\in Z((i,j)\in E)$①),可以如下构造 $GPP(V,E,c,w)$ 的算法:(1) 根据定义 3 可知,由 $GPP(V,E,c,w)$ 可在 $O(n^2)$ 时间内构造偏移实例 $GPP(V,\hat{E},c,\hat{w})$;(2) 由于 $GPP(V,\hat{E},c,\hat{w})$ 也属于 GPP 问题,按照假设,可利用算法 Λ 在 $O(\Diamond)$ 得到其骨架;(3) 因 $GPP(V,\hat{E},c,\hat{w})$ 仅有一个全局最优解(定理 1),故此其骨架就是全局最优解的割,再由算法 1(割的可行解构造算法)可在 $O(kn+n)$ 内得到可行解,而根据定理 1 可知,该可行解就是全局最优解. 因此,通过步(1)~(3)可以在 $O(\Diamond)+O(kn+n)+O(n^2)$ 时间内得到原实例的一个全局最优解,由此推出:GPP 问题是多项式时间内可解的. 而根据计算复杂性理论可知:除非 P=NP,否则 NP-难解问题不存在多项式时间的完全算法,矛盾. 故假设不成立,原命题得证. 证毕.

4 基于偏移实例的 IBS 算法

本节应用第 3 节的结论,提出基于偏移实例的 IBS 算法(BI-IBS). IBS 算法是由 Zou 等[17]给出的求解 GPP 问题的元启发算法,其基本原理是利用局部最优解的交集来模拟骨架,并以此实现对于 GPP 实例的归约. 由于归约后新实例规模缩小,故此可以用现有的各种启发式算法高效地求解. 然而该方法存在一定缺陷:根据定义 4 可知,GPP 实例 $GPP(V,E,c,w)$ 的全局最优解越多,骨架的规模就越小,此时归约的效果也越差. 一种极端的例子是:当 $GPP(V,E,c,w)$ 满足 $E=\{(i,j)\mid i,j\in V,i\neq j\}$,$c_i=1(i\in V)$,$w_{ij}=1((i,j)\in E)$ 时,$bone(V,E,c,w)=\varnothing$,此时骨架在算法中就不起任何作用了. 因此,寻求提高骨架规模的方法对于基于骨架的算法有重要意义. 而本文在第 3 节中已经证明:偏移 GPP 实例具有唯一的全局最优解(定理 1),且该解也是原实例的全局最优解(引理 2). 因此,可以利用偏移实例对 IBS 算法进行改进.

由于偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$ 的特殊性,下面先给出专门针对其的归约算法(算法 2). 算法 2 的主要工作包括:(1) 利用给定的局部最优解,构造新实例的顶点集合,方法是:若原实例的顶点 i,j 在所有给定局部最优解中均属于同一子集(就偏移 GPP 实例而言,即要求满足 $(i,j)\notin\hat{C}(s_1)\cup\hat{C}(s_2)\cup\cdots\cup\hat{C}(s_h)$),则在新实例中合并 i,j;(2) 在新的顶点集(新的顶点可能对应原实例的若干顶点)上构造新的边集合和边权值.

算法 2. 偏移 GPP 实例归约.

输入:GPP 偏移实例 $GPP(V,\hat{E},c,\hat{w})$,h 个局部最优解 $s_1,s_2,\cdots,s_h$

输出:新的实例 $GPP(V',E',c',w')$

Begin

1. $V^0=V$, $C_s=\hat{C}(s_1)\cup\hat{C}(s_2)\cup\cdots\cup\hat{C}(s_h)$
2. $V'=\varnothing$, $V''=\varnothing$, $E'=\varnothing$, $t=0$;
3. while $V^0\neq\varnothing$ do //构造新实例的顶点集合
   - 3.1. $t=t+1$;
   - 3.2. 取 $i\in V^0$, $V^0=V^0-i$, $V''_t=\{i\}$;
   - 3.3. for each $j\in V^0$ do
     if $(i,j)\notin C_s$ then $V''_t=V''_t\cup\{j\}$, $V^0=V^0-j$;
   - 3.4. $V'=V'\cup\{t\}$;
   - 3.5. $V''=V''\cup\{V''_t\}$; //每一个顶点集合 V''_t 对应 //新实例的一个顶点
   - 3.6. $c'_t=\sum_{p\in V''_t}c_p$; //新实例的顶点权值
4. for $i=1$ to t do //构造新实例的边集合和边权值
   for $j=i+1$ to t do
   - 4.1. $E'=E'\cup\{(i,j)\}$;
   - 4.2. $w'_{ij}=\sum_{p\in V''_i,q\in V''_j}\hat{w}_{pq}$;
5. return $GPP(V',E',c',w')$

End

算法 2 的时间复杂度分析如下:算法步 1 的时间复杂度为 $O(hn^2)$. 步 2 的时间复杂度为 $O(1)$. 算法步 3.1、3.2、3.4 和步 3.5 的时间复杂度均为

① 如果存在边的权值为小数的情况,可以将所有的边的权值放大相同倍数,得到新的权值 w'_{ij} 均为整数,此时新实例 $GPP(V,E,c,w')$ 的解与原实例 $GPP(V,E,c,w)$ 相同.

$O(1)$. 在最坏情况下(即每一个顶点集合 V_t'' 均仅包含一个顶点时),步 3.3 每次执行所需时间为 $O(n-t)$(其中,$t=1,2,\cdots,n$). 注意到整个步 3 执行过程中原实例的每一顶点权值恰好在步 3.6 用到一次,故步 3.6 在整个算法运行过程中所需时间为 $O(n)$. 整个步 3 的时间复杂度为 $O(n^2)$. 算法步 4.1 的时间复杂度为 $O(1)$,注意到整个步 4 执行过程中原实例的每一条边权值恰好在步 4.2 用到一次,故此步 4 的时间复杂度为 $O(n(n-1)/2)+O(t(t-1))=O(n^2)$. 故此整个算法 2 的时间复杂度为 $O(hn^2)$.

在给出偏移 GPP 实例归约算法的基础上,算法 3 给出了 BI-IBS 的算法描述. 算法 3 的时间复杂度分析如下:步 1 的时间复杂度为 $O(n^2)$;记算法 A 求解 $GPP(V,\hat{E},c,\hat{w})$ 的时间复杂度为 $O(*)$,则步 2 的时间复杂度为 $hO(*)$;步 3 使用的是算法 2,时间复杂度为 $O(hn^2)$;步 4 中 A 求解 $GPP(V',E',c',w')$ 的时间复杂度记为 $O(\circ)$;步 5 的时间复杂度为 $O(n)$. 综上可知,算法 3 的时间复杂度为 $O(hn^2)+hO(*)+O(\circ)$.

算法 3. BI-IBS.

输入:GPP 实例 $GPP(V,E,c,w)$,h,现有求解 GPP 算法 A

输出:解 s

Begin

1. 构造偏移 GPP 实例 $GPP(V,\hat{E},c,\hat{w})$;
2. 调用 A 算法 h 次,获得 $GPP(V,\hat{E},c,\hat{w})$ 的 h 个局部最优解 $s_1,s_2,\cdots,s_h$;
3. 利用偏移实例归约算法对 $GPP(V,\hat{E},c,\hat{w})$ 进行归约,得到新的规模更小实例 $GPP(V',E',c',w')$;
4. 利用 A 算法求解 $GPP(V',E',c',w')$,得到解 s';
5. 将 s' 转化为 $GPP(V,\hat{E},c,\hat{w})$ 的解 s;

End

5 实验及分析

本文以 Kernighan-Lin 算法作为从属算法,用 C++语言实现了元启发算法 BI-IBS(记为 BI-IBS-KL)、IBS(记为 IBS-KL)和 Multilevel(记为 Multilevel-KL). 实验的测试实例 $G_{n.p}$ 采用研究人员普遍使用的随机方法[5,8]生成,其中 n 代表顶点个数,p 代表任意两个顶点间存在边的概率,顶点和边的权值均设为 1,这样顶点的度(degree)的期望值为 $p(n-1)$. 实验参数选取 $k=2$,局部最优解的个数 $h=3$,每个测试实例运行 20 次. 实验环境是 P4 2.4GHz 微处理机,512MB 内存,Redhat Linux 9.0 操作系统. 表 1 给出了算法 BI-IBS-KL、IBS-KL 和 Multilevel-KL 在测试实例上的实验结果,其中 w_{best}、w_{aver} 分别代表 20 次运行结果的目标函数值最优值、平均值.

表 1 算法 BI-IBS-KL、IBS-KL、Multilevel-KL 在 $G_{n.p}$ 实例上的实验结果

实例	BI-IBS-KL 上的结果		IBS-KL 上的结果		Multilevel-KL 上的结果	
	w_{best}	w_{aver}	w_{best}	w_{aver}	w_{best}	w_{aver}
$G_{200.005}$	2	2	2	2	2	2
$G_{200.01}$	4	4	4	4	4	4
$G_{200.02}$	11	11.5	11	11.7	11	11.9
$G_{200.03}$	35	36.5	35	37.2	36	38.1
$G_{200.04}$	87	89.1	87	88.6	89	91.2
$G_{200.05}$	172	174.4	174	176.9	176	179.3
$G_{500.005}$	49	50.9	49	51.1	49	51.2
$G_{500.01}$	220	222.3	223	225.3	222	224.4
$G_{500.02}$	625	627.2	629	632.1	631	634.9
$G_{500.03}$	1127	1128.9	1129	1133.4	1131	1135.7
$G_{500.04}$	1735	1739.6	1742	1748.7	1745	1750
$G_{500.05}$	2589	2598.4	2591	2602.3	2594	2606.6
$G_{1000.005}$	443	448.4	446	452.1	447	452.8
$G_{1000.01}$	1375	1381.6	1381	1389.4	1381	1390.2
$G_{1000.02}$	3410	3418.5	3421	3431.6	3423	3433.6
$G_{1000.03}$	7869	7879.7	7890	7907.8	7892	7911.3
$G_{1000.04}$	11201	11279.1	11320	11420.6	11345	11432.8
$G_{1000.05}$	18930	18983.2	18990	19030.4	19012	19087.9

从表 1 中可以发现,算法 BI-IBS-KL、IBS-KL、Multilevel-KL 在小规模实例 $G_{200.005}$、$G_{200.01}$ 上的结果相同,其原因是对于稀疏图上的小规模实例,从属算法 Kernighan-Lin 本身的求解能力很强. BI-IBS-KL 在绝大多数实例上的平均值优于 IBS-KL,在 6 个实例上与 IBS-KL 的最优值相同,在其余实例上均不同程度地优于 IBS-KL. BI-IBS-KL 在 16 个实例上的平均值优于 Multilevel-KL,在 4 个实例上与 Multilevel-KL 的最优值相同,在其余 14 个实例上均优于 Multilevel-KL. 因此,算法 BI-IBS-KL 的性能在整体上优于 IBS-KL、Multilevel-KL. 同时,随着实例规模(包括顶点和边的数目)的增加,所有算法的平均值与最优值的差距逐渐增加,反映出 GPP 实例的难度随着规模增加而迅速提高.

6 结 论

本文首次对 GPP 问题的骨架进行了理论分析,通过构造具有唯一全局最优解的偏移 GPP 实例,本文不仅证明了获取 GPP 问题的骨架属于 NP-难解问题,而且有效地提高了骨架的规模. 利用偏移 GPP 实例构造方法,改进了目前求解 GPP 问题最

好算法之一的 IBS 算法. 论文工作为 NP-难解问题骨架的理论分析及利用骨架来进行启发式算法设计提供了一种可行的途径.

参 考 文 献

[1] Xu Da-Chuan, Han Ji-Ye, Du Dong-Lei. Improved approximation algorithm for graph partition problem. Acta Mathematicae Applicatae Sinica, 2005, 28(4): 587-597(in Chinese) (徐大川,韩继业,杜东雷. 关于图划分问题的改进的近似算法. 应用数学学报, 2005, 28(4): 587-597)

[2] Han Q, Ye Y, Zhang J. An improved rounding method and semidefinite relaxation for graph partition. Mathematical Programming, 2002, 92(3): 509-535

[3] Kernighan B, Lin S. An efficient heuristic procedure for partitioning graphs. Bell Systems Technical Journal, 1970, 49 (2): 291-307

[4] Batti R, Bertossi A. Differential greedy for the 0-1 equicut problem//Proceedings of the DIMACS Workshop on Network Design: Connectivity and Facilities Location. Nassau Inn, Princeton, NJ, 1997: 3-23

[5] Johnson D S, Aragon C R, McGeoch L A. Schevon C. Optimization by simulated annealing: Part I, graph partitioning. Operations Research, 1989, 37: 865-892

[6] Batti R, Bertossi A. Greedy, prohibition, and reactive heuristics for graph-partitioning. IEEE Transactions on Computers, 1999, 48(4): 361-385

[7] Banos R, Gil C, Ortega J, Montoya M G. A parallel evolutionary algorithm for circuit partitioning//Proceedings of the 11th Euromicro Workshop on Parallel and Distributed Processing. Genova, Italy, 2003: 365-371

[8] Merz P, Freisleben B. Fitness landscapes and Memetic algorithms and greedy operators for graph bi-partitioning. Evolutionary Computation, 2000, 8(1): 61-91

[9] Soper A J, Walshaw C, Cross M. A combined evolutionary search and Multilevel optimisation approach to graph-partitioning. Journal of Global Optimization, 2004, 29(2): 225-241

[10] Slaney J, Walsh T. Backbones in optimization and approximation//Proceedings of the 17th International Joint Conference on Artificial Intelligence. Seattle, 2001: 254-259

[11] Monasson R, Zecchina R, Kirkpatrick S, Selman B, Troyansky L. Determining computational complexity from characteristic ‘phase transition’. Nature, 1998, 400(2): 133-137

[12] Zhang W X. Phase transition and backbones of the asymmetric traveling salesman problem. Journal of Artificial Intelligence Research, 2004, 21(1): 471-497

[13] Schneider J. Searching for backbones — A high-performance parallel algorithm for solving combinatorial optimization problems. Future Generation Computer Systems, 2003, 19 (1): 121-131

[14] Zhang W X. Configuration landscape analysis and backbone guided local search: Part I: Satisifiability and maximum satisfiability. Artificial Intelligence, 2004, 158(1): 1-26

[15] Dubois O, Seymour P. A backbone-search heuristic for efficient solving of hard 3-SAT formula//Proceedings of the 17th International Joint Conference on Artificial Intelligence, Seattle, 2001: 248-253

[16] Zou Peng, Zhou Zhi, Chen Guo-Liang, Jiang He, Gu Jun. Approximate-backbone guided fast ant algorithms to QAP. Journal of Software, 2005, 16(10): 1691-1698(in Chinese) (邹鹏,周智,陈国良,江贺,顾钧. 求解 QAP 问题的近似骨架导向快速蚁群算法(英). 软件学报, 2005, 16(10): 1691-1698)

[17] Zou P, Zhou Z, Wan Y Y, Chen G L, Gu J. New meta-heuristic for combinatorial optimization problems: Intersection based scaling. Journal of Computer Science and Technology, 2004, 19(6): 740-751

[18] Kilby P, Slaney J, Walsh T. The backbone of the traveling salesperson//Proceedings of the 19th International Joint Conference on Artificial Intelligence. Edinburgh, Scotland, UK, 2005: 175-181

JIANG He, born in 1980, Ph. D., associate professor. His research interests include intelligent computation, data mining.

QIU Tie, born in 1980, Ph. D. candidate. His research interests focus on data mining.

Background

This project is supported by the National Natural Science Foundation of China (60805024), the National Research Foundation for the Doctoral Program of Higher Education of China (20070141020). It focuses on the efficient heuristics for classic NP-Hard problems (such as SAT, TSP, Bin-Packing and GPP).

The group has presented some fast heuristics and theoretic results of NP-Hard problems and published many papers on Journal of Combinatorial Optimization, Discrete Mathematics, Information Processing Letters, Journal of Computer Science and Technology, Journal of Advanced Software, Chinese Journal of Computers (in Chinese), Journal of Software (in Chinese), etc. And being a part of this project, this paper mainly focuses on the heuristic design of the graph partitioning problem.