ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, No.9, September 2009, pp.2307−2319
doi: 10.3724/SP.J.1001.2009.03593
© by *Institute of Software*, *the Chinese Academy of Sciences*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

参数计算中核心化技术及其应用*

李绍华[1,2], 王建新[1+], 冯启龙[1], 陈建二[1]

[1](中南大学 信息科学与工程学院,湖南 长沙 410083)
[2](广东商学院 信息学院,广东 广州 510320)

Kernelization Techniques and Its Applications to Parameterized Computation

LI Shao-Hua[1,2], WANG Jian-Xin[1+], FENG Qi-Long[1], CHEN Jian-Er[1]

[1](School of Information Science and Engineering, Central South University, Changsha 410083, China)
[2](Information School, Guangdong University of Commercial Studies, Guangzhou 510320, China)
+ Corresponding author: E-mail: jxwang@csu.edu.cn

Li SH, Wang JX, Feng QL, Chen JE. Kernelization techniques and its applications to parameterized computation. *Journal of Software*, 2009,20(9):2307−2319. http://www.jos.org.cn/1000-9825/3593.htm

Abstract: According to parameterized complexity theory, a decidable parameterized problem is fixed-parameter tractable if and only if it can be kernelized. Kernelization is the most widely applied and effective technique in the parameterized algorithm design. It is one of the hottest issues in parameterized complexity theory. This paper firstly introduces four main kernelization techniques, which are compared and analyzed with practical examples. Then it discusses how to apply these techniques to parameterized problems, such as covering problems, packing problems and cutting problems. Finally, the paper gives the future research directions about kernelization, especially the new possible kernelization technique and the kernel optimization of several FPT problems.
Key words: kernelization; crown decomposition; extremal induction; randomized algorithm; fixed-parameter tractable

摘 要: 在参数计算与复杂性理论中,一个参数问题是固定参数可解的问题当且仅当该问题是可核心化的.核心化技术是参数化算法设计中应用最为广泛、有效的技术,是参数理论中的一个研究热点.通过实例分析对比了最主要的4种核心化技术的基本思想、应用特点和方法,总结了核心化技术在cover类、packing类和cut类等几个重要领域中的应用成果,展望核心化技术的进一步研究方向并加以分析讨论,针对核心化新技术研究和某些热点问题,提出了可能采取的核心优化方法和思路.
关键词: 核心化;皇冠分解;极值归纳;随机算法;固定参数可解
中图法分类号: TP301 **文献标识码**: A

* Supported by the National Basic Research Program of China under Grant No.2008CB317107 (国家重点基础研究发展计划(973)); the National Natural Science Foundation of China under Grant Nos.60433020, 60773111 (国家自然科学基金); the Program for New Century Excellent Talents in University of China under Grant No.NCET-05-0683 (新世纪优秀人才支持计划); the Program for Cheung Kong Scholars and Innovative Research Team in University of China under Grant No.IRT0661 (长江学者和创新团队发展计划)

Received 2008-11-06; Revised 2009-01-14; Accepted 2009-02-16

© 中国科学院软件研究所 http://www.jos.org.cn

近年发展起来的参数计算与复杂性理论[1,2]是理论计算机科学的一个重要分支.该理论的目的之一是针对 NP 难解的计算问题,通过应用参数理论的算法设计技术,得到称为固定参数可解的精确解,使得在 NP 完全理论中只存在指数级时间复杂性的精确解问题,在参数计算理论中找到实际可行的精确解.达到这一目的的前提是基于这样一个事实:许多 NP 难解问题在实际应用中存在一个小参数.例如,顶点覆盖问题是一个经典的 NP 难问题.通过应用参数理论的核心化技术,设计出 $O(kn+1.286^k)$的精确算法[3],这一算法成果已被瑞士的计算化学生物组的 Darwin 项目所采用,以解决构建分子进化树过程中基因序列之间的冲突问题.冲突问题发生在构建分子进化树过程中所采用的多序列相似性的比对.进行多序列比对时,首先需要插入空位(gaps)来对齐各序列.当对 n 个序列插入多个空位后,若序列集 S_1 均有空位 g_i,序列集 S_2 均有空位 g_j,而 $S_1\cup S_2=\varnothing$,便称空位之间产生了冲突,这在分子进化意义上是不允许的,我们必须把尽可能少的空位从序列中排除掉,使得保留下来的空位不存在冲突.文献[4]通过把空位看成图的一个顶点,顶点之间的边表示两个空位有冲突,则删除图中的 k 个顶点使得所有的冲突边均被删除,这个问题等价于 k 顶点覆盖问题.由于在实际的序列比对中,空位的数量不多于 60 个[1,4],因此,时间复杂度为 $O(kn+1.286^k)$的参数化顶点覆盖精确算法是实际可行的.在传统的 NP 完全理论中[5],求解诸如点覆盖这类 NP 难解问题的算法一般应用启发式算法、近似算法和随机算法等技术,这类方法均无法或难以得到可行的精确解.

当我们用参数理论来求解 NP 难问题时,首先要将该问题转化为参数化问题.参数化问题 Q 是一个用二元组(x,k)表示的判定问题,其中,x 代表一个具体的问题,k 是一个非负整数,称为参数.然后应用参数理论的算法技术对参数化问题 Q 加以求解.若我们能够设计一种参数化求解算法,其时间复杂度形如 $O(f(k)|x|^c)$,则我们把该问题 Q 称为固定参数可解的问题,简称 FPT 问题.若参数 $k<<|x|$,则称是小参数问题.对于具有小参数特征的 NP 完全的参数问题 Q,若能设计出 FPT 算法,该问题就是参数易解的问题了.

FPT 算法设计技术的研究一直是参数理论中的研究热点.设计 FPT 算法的技术主要有分支搜索(branching search)技术[6−12]、核心化(kernelization)技术[2,10]及递归压缩(iterative compression)技术[7]等(如图 1 所示).2005 年,Chen 在文献[1]中系统地介绍了参数计算与复杂性理论中常用的 FPT 算法设计技术、下界技术、参数近似理论等的基本原理和基本概念.由于一个参数化问题是 FPT 问题当且仅当该问题是可核心化的[1],因此,在参数算法设计中,核心化技术应用最为广泛.本文着重展开讨论局部简化、皇冠分解、极值归约和随机算法这 4 种常用的核心化技术的基本思想和技术方法,并加以对比分析,同时总结了核心化技术在一些重要领域中的应用成果.

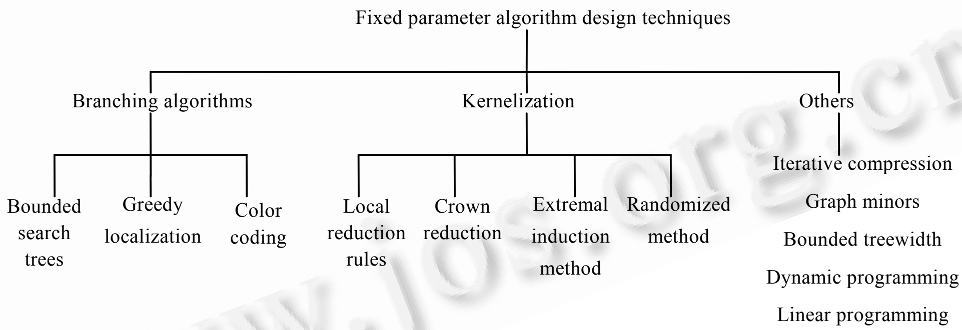

Fig.1 Main techniques for designing FPT algorithms
图 1 FPT 算法设计中的主要技术

本文第 1 节简要介绍核心化的相关定义和应用核心化技术求解参数问题的基本思路.第 2 节分析比较 4 种常用的核心化技术的基本思想和方法.第 3 节介绍核心化技术在几个重要领域中的应用.第 4 节给出全文的总结,并指出核心化技术的进一步研究方向.

© 中国科学院软件研究所 http://www.jos.org.cn

1 核心化思想

核心化是指如果存在一个多项式时间算法 K 和一个递归函数 g,使得对于参数问题 Q 的任意一个实例 (x,k),应用算法 K 将 Q 转化为一个新实例 (x',k'),使得 $|x'|\le g(k)$ 和 $k'\le k$,并满足当且仅当 (x',k') 是参数问题 Q 的一个真实例,(x,k) 也是 Q 的一个真实例,则我们说参数问题 Q 是可核心化的,算法 K 称为核心化算法,新实例 (x',k') 称为核心化后的问题核.从以上定义我们可以看出,用核心化技术求解参数化问题的基本步骤为:① 对实例加以简化:设计一个核心化算法 K,使得原问题 Q 的一个实例 (x,k) 转化为另一个实例 (x',k'),且 $|x'|\le g(k)$ 和 $k'\le k$;② 对实例 (x',k') 加以求解,从而得到原问题 Q 本身的解.核心化后问题规模极大地降低.核的大小可能是参数 k 的线性式(线性核),也可能是参数 k 的多项式(多项式核),甚至是参数 k 的非多项式.

核心化的基本思想就是应用核心化算法 K 中的简化规则,对原问题实例进行预处理,进而得到问题的足够小的核.由于问题规模极大地降低,以至于我们对简化后的实例施以穷举之类的 brute-force 算法,便可得到可行解了.核心化技术主要有 4 种:局部简化技术、皇冠分解技术、极值归纳技术和随机算法技术.从本质上来说,任何一种核心化技术均是针对具体的问题,从不同的角度、采用不同的处理方法,设计出高效的简化规则,以达到降低问题实例的大小,获得更小的核的目的.下一节将对这 4 种技术的基本思想和方法特点加以分析讨论.

2 核心化技术

2.1 局部简化技术

局部简化技术是通过设计出一系列可以在多项式时间内完成的简化规则,然后应用这些简化规则对问题加以简化.若经简化后的问题结构具有固定的大小,则该类简化规则称为局部简化规则.对局部简化规则须加以证明,以使得简化后的问题实例有正确解,当且仅当原问题实例也有正确解.几乎所有的 FPT 算法都会应用该技术,对问题进行预处理.如果一个简化规则集合能将一个问题的任何一个实例简化为核,我们就说该简化规则集合是一个完全简化规则集合.下面以参数化点覆盖问题为例来说明局部简化技术的应用.

参数化点覆盖问题(k-VC)指的是,给定一个图 $G(V,E)$ 和一个整数 k,$|V|=n$,判断图 G 是否存在一个至多 k 个点的点覆盖.k-VC 问题的简化规则集合包括低度点的简化规则和高度点的简化规则.低度点的简化规则有以下 4 个:① 度为 0 的点:直接删除.② 度为 1 的点:将其邻居 $N(v)$ 加入到点覆盖集中.③ 度为 2 且两个邻居之间存在边的点:删除这 3 个点以及所有的关联边,并将这两个邻居加入到最小点覆盖中.④ 度为 2 且两个邻居之间不存在边的点:进行 2 度点的折叠.设 v 邻居为 u 和 w 且 u 和 w 不相邻.用一个新的点 v' 来替代这 3 个点,将 u 和 w 的所有邻居作为 v' 的邻居.k-VC 问题的高度点简化规则如下:如果 G 有一个度大于 k 的点 v,则将其点加入到点覆盖集中,则 (G,k) 可简化为 $(G-v,k-1)$.这是因为度大于 k 的点必定在任何一个 k-VC 的覆盖集 C 中,否则将需要多于 k 个点才能覆盖到点 v 的所有邻接边.

对实例 (G,k) 重复使用上述各规则,直到无法再加以简化为止.设 $(G(V,E),k)$ 最终核化为 $(G'(V',E'),k')$, $k'\le k$ .C 是 (G',k') 的点覆盖, $|C|=k'$,则 $F=V'-C$ 便是 G' 的独立集.由于 F 中任一点的度都大于 2,故 F 中的任一点在 C 中至少有 3 个邻居,即从 F 中引出的边数 $\ge 3(n'-k')$,而 C 中任一点的度 $\le k'$,因此 $3(n'-k')\le |F|\le k'^2$, $n'\le k'^2/3+k'$.即实例 (G,k) 核化后可得到核的大小为 $k^2/3+k$ 的等价实例 (G',k').

从上例可以看出,局部简化技术的基本思路是:分析具体问题的特性,得出完全简化规则集合,应用该规则集进行预处理,直至该问题没有简化规则可应用的结构,从而得到与问题规模 n 无关的核 $f(k)$,核心化后,再应用诸如分支搜索树技术得到问题的 FPT 解.

2.2 皇冠分解技术

2003 年,Fellows[13]首次提出皇冠分解的概念.如果 I 是 $G=(V,E)$ 的一个非空独立集,$N(I)$ 是 I 的邻居,且 $N(I)$ 是二分图 $G(N(I)\cup I)$ 的最小点覆盖,则称 $N(I)\cup I$ 为皇冠结构.皇冠分解是将图 $G=(V,E)$ 的顶点集 V 划分成 3 个点集 H,I,R,其中,H 是图 G 的一个分割点集,使得图 G 中属于 I 集的点和属于 R 集的点之间没有边相连;I 是图 G

的一个独立集,将I分成两个子集I_u和I_m,I_m和H之间存在一个完美匹配.图2是一个$|H|=3$的皇冠分解结构示例,3个椭圆标识出图G的顶点集V划分成3个顶点子集H,I,R,其中,$I=I_m\cup I_u$,并用粗实线表示出I_m与H集构成的一个完美匹配.

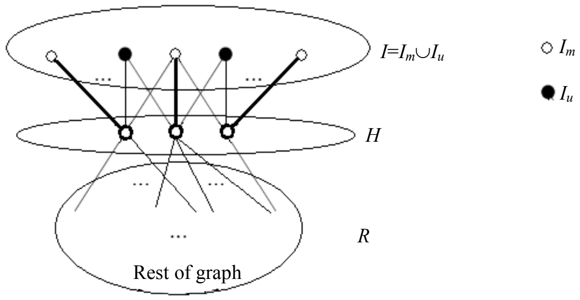

Fig.2 Sample crown decomposition of width 3

图2 皇冠分解示例:$|H|=3$

皇冠分解的基本思想是:在图$G=(V,E)$中找到皇冠结构$I\cup H$,删除该皇冠以降低图规模,重复进行这类查找和删除皇冠的操作,直到图中不再有任何皇冠为止.在实际应用中,一般采用贪心算法找到图G的一个极大匹配M,然后根据具体问题采用不同的策略进行点集的皇冠划分,直至无法继续找到皇冠.对于皇冠分解的时间复杂性,文献[14]的引理5表明,只要在图G中能找出一个独立集I,且$|I|$不比其邻接点少,则可以在多项式时间内找到图G的一个皇冠分解.下面我们依然以k-VC问题为例来说明皇冠分解技术的应用.

如果G有一个皇冠$I\cup H$,则G中存在一个包含H中所有的点但不包含任何I中的点的最小点覆盖.找到图G的一个皇冠后,把H部分加入到点覆盖的集合中,然后将$I\cup H$包含的点及其所有在G中的关联边删除,形成一个新图$G'=G[V(G)-V(I\cup H)]$,则G'便是对G进行一次皇冠分解操作后的核.重复应用皇冠分解直到不能发现新的皇冠为止,问题核便越来越小.为找到图G的皇冠结构,可先找到图G的一个极大匹配M_1,而未被匹配的点的集合用O表示;再在O和$N(O)$之间找到一个辅助最大匹配M_2,O中未被M_2匹配的点的集合用I_0表示;然后从$n=0$开始,重复地执行$H_n=N(I_n)$和$I_{n+1}=I_n\cup N_{M_2}(H_n)$,直至$I_{n-1}=I_n$为止,则$(I_n,H_n)$便是一个皇冠.

分析以上算法可以看出,如果匹配M_1和M_2的大小有任何一个大于k,则图中必不存在这样的点覆盖;如果匹配M_1和M_2的大小都小于或等于k,则M_1至多包含$2k$个顶点,所以O至少包含$n-2k$个顶点.由于M_2小于或等于k,因此O中至多有k个点与M_2匹配,既O中至少有$n-3k$个点未与M_2匹配,这些未匹配的点均为I集中的点,所以,图G最多有$3k$个点不在皇冠中,即核为$3k$.以上各步骤用时为$O(n^{5/2})$,详细证明可参见文献[15].

皇冠分解技术得到了广泛的应用,如p_2-Packing($7k$核)[16]和顶点不相交三角形packing($O(k^3)$核)[17]等问题.皇冠分解基于一些不能确定固定大小结构的规约规则,所以皇冠分解的核化效果比局部简化技术更加明显.

2.3 极值归纳技术

极值归纳(coordinated kernel)技术由Prieto[18]于2000年提出.该技术适于求解极大化的参数问题,且要求问题关于参数是单调的,即如果实例(G,k)是一个no实例,则对于所有的$k'>k$,(G,k')也是一个no实例.

极值归纳技术的基本思想是:① 先假设(G,k)是yes实例,而$(G,k+1)$是no实例.② 若这一假设成立,则该参数问题一定具有某些性质特征,依据性质特征集P可以得到简化规则集R,并把R应用到(G,k)中直到不能再简化.③ 对于简化后得到的图G',若其顶点集$V(G')$的大小不大于$f(k)$,可得出以下结论:若(G,k)是yes实例而$(G,k+1)$是no实例,则$|V(G)|\le f(k)$一定成立.这一结论称为界值引理.④ 反之,若$|V(G)|>f(k)$,与“顶点集$V(G')$的大小不大于$f(k)$”矛盾,则说明以上假设不成立,即$(G,k+1)$必然也是yes实例.这一结论称为核化引理.

当我们应用极值归纳技术来求解实际问题时,由问题实例(G,k')出发递归计算$(G,k'+1)$.因为问题是参数单调的,所以如果(G,k')是no实例,则$(G,k'+1)$也是no实例.若“(G,k)是yes实例,而$(G,k+1)$是no实例”,则G应具有

性质特征集 P.如果 G 不具有这类性质特征,则由于 (G,k') 是 yes 实例,因此 $(G,k'+1)$ 也是 yes 实例了.如此一直判断下去,对于极大化参数问题,我们一定会遇到一个 $(G,k+1)$ 是 no 实例的.于是便可根据界值引理得出以下结论: $|V(G)|\le f(k)$,从而得到了问题核的大小为 $f(k)$.

下面以 k-边不相交三角形 packing 问题(简称 k_3-packing 问题:给定一个图 $G=(V,E)$ 和正整数 k,问 G 中是否存在至少 k 个边不相交的三角形)为例,介绍极值归纳技术的具体应用方法.

k_3-packing 问题是关于参数 k 单调的,即若 G 不存在 k 个边不相交的三角形,则一定不存在 $k+1$ 个边不相交的三角形.文献[14]首先给出 3 个局部简化规则和皇冠分解规则.设经过以上规则化简后,图 G 化简为 G'.我们把 G' 分为 P 和 O 两部分,其中 P 是一个大小为 k 的边不相交三角形 packing,显然 $|P|\le 3k$,O 是剩余的部分.再把 P 分成 3 部分: P_0(三角形中的边没有被 O 中的点支配)、P_1(三角形的边中有一条边被 O 中的点支配)、P_2(三角形的三条边都被 O 中的点支配).假设 (G,k) 是真实例而 $(G,k+1)$ 是假实例,对 (G,k) 应用局部简化规则和皇冠分解规则后,再结合 k_3-packing 的特点,便可分析得出假设成立的数个性质特征[14],进而证明 $|O|$ 一定小于等于 k,否则假设便不成立.于是,我们可以得出以下的两个结论:① 如果 $G=(V,E)$ 已简化,且 $|V|>4k$,则 G 是 k_3-packing 的真实例. ② 若 G 已简化且 (G,k) 是问题的真实例,而 $(G,k+1)$ 是问题的假实例,则 $|V|\le 4k$.从以上两个结论可以看出,k_3-packing 问题核为 $4k$.

极值归纳技术在许多问题上得到了成功的应用,如: Cluster Editing(4k 核)[19]、k-Internal Spanning Tree($O(k^3)$ 核)[20]以及 $K_{1,s}$-packing($O(k^2)$ 核)[21]等问题.具体应用极值归纳技术进行核心化时,主要着眼点是得出界值引理和核化引理,然后从实例 $(G,0)$ 出发加以递推,因为对于最大化问题来说,$(G,0)$ 显然是一个真实例.

2.4 随机算法技术

随机方法是一种传统的算法设计技术.在 FPT 算法设计中,随机方法有很多成功的应用(如 k-path,3-set packing,3-D matching 等),Chen 等人[22]首次将随机方法应用于无权 set splitting 问题的核化过程中,得到 $2k$ 的问题核,方法简单、有效.

给定 set splitting 问题的一个实例 (X,F,k),其中 F 含有 n 个集合,每个集合的元素都来自符号集 X.set splitting 问题的求解目标是对 X 进行划分,使得 F 中至少有 k 个集合被分割.对 X 进行划分就是把 X 分成两个子集 (X_1,X_2),使得 $X_1\cup X_2=X$ 且 $X_1\cap X_2=\varnothing$.对于 F 中的一个集合 S,如果 S 与 X_1 和 X_2 的交集都不为空,就称 S 被划分 (X_1,X_2) 所分割.

Set splitting 是 NP 难问题[5],在近似理论中,该问题是 APX 完全的,Zhang 等人[23]提出一种改进的多项式时间近似算法,近似率为 0.7499.在参数化算法设计方面,Dehne 等人[24]首次提出了 Set splitting 的 FPT 算法,时间复杂度为 $O(72^k)$.Lokshtanov[25]提出一种 $O(2.65^k)$ 的 FPT 算法.目前最好的算法是 Chen 在文献[22]中提出的基于随机方法的 FPT 核化算法,问题核的大小为 $2k$,运行时间为 $O(2^k)$,该算法基本思路如下:

文献[24]首先给出两个简化规则,并应用到问题实例 (X,F,k) 中,便可删除只含有 1 个元素的集合和包含至少 k 个元素的集合,从而得到问题的新实例 (X',F',k').并证明了若 $|F'|\ge 2k$,则 (X,F,k) 为真实例,否则 $|F'|<2k,k'<k$.简化规则 2 表明 F' 中的每个集合最多包含了 $k-1$ 个元素,因此 $|X'|<2k^2$,该过程可在 $O(n+2k^2)$ 时间内完成.

概率上 X 中的每个元素以同样的 1/2 概率被划分到 X_1 和 X_2 中.集合中的元素越多,被分割的概率就越大.假设集合 F 中含有 i 个元素的集合个数为 $m_i(1\le i\le k-1)$,m_k 为至少 k 个元素的集合个数,$F_{<k}$ 是 F 中所有含有元素数不超过 k 的集合.对于 $F_{<k}$ 中的任一元素个数为 i 的集合 S,当对 X 进行划分时,S 被分割的概率为 $Pr(S)=\dfrac{2^i-2}{2^i}$.引入随机变量 Xs,若 S 被分割,则 $Xs=1$,否则 $Xs=0$.$F_{<k}$ 中被分割的集合数的期望值为

$$E\left(\sum_{S\in F_{<k}} X_s\right)=\sum_{i=1}^{k-1}\sum_{|s|=i}E(S)=\sum_{i=1}^{k-1}\frac{2^i-2}{2^i}m_i,$$

因此,存在 X 的一种划分使得 $F_{<k}$ 中至少有 $\sum_{i=1}^{k-1}\frac{2^i-2}{2^i}m_i$ 个元素被分割.如果 $\sum_{i=1}^{k-1}\frac{2^i-2}{2^i}m_i\ge k-m_k$,则 X 必定存在一种

© 中国科学院软件研究所 http://www.jos.org.cn

划分使得 $F_{<k}$ 中至少有 $k-m_k$ 个集合被分割,否则 $\sum_{i=1}^{k-1}\frac{2^i-2}{2^i}m_i+m_k<k$,即 $|F'|=\sum_{i=1}^{k-1}m_i<2k-\sum_{i=3}^{k-1}\frac{2^{i-1}-2}{2^{i-1}}m_i-2m_k$.因此,问题核的大小为 $|F'|<2k$,找到核的时间复杂度为 $O(n+2k^2)$.

生物信息学中有许多实际问题,其元素具有不同程度的随机性和不确定性,运用随机方法对这类问题加以核化,是值得进一步研究的方向.

3 核心化技术的应用

核心化技术已在许多领域得到了实际应用.限于篇幅,本节主要从 Cover 类、Packing 类和 Cut 类这 3 类问题来反映核心化技术的应用.

3.1 Cover类问题

覆盖问题是图论中的经典问题,应用非常广泛.我们从 Vertex Cover、Clique Cover 和 Hyperplane Cover 这三个覆盖问题来阐述核心化技术的成功应用.

作为图论中经典的 NP 完全问题之一,许多实际问题均可以模型化为 k-VC 问题,对该问题的算法研究至今都是一个热点.文献[26]对 k-VC 问题的研究现状加以讨论.1993 年,Buss 等人[27]基于 Buss 的高度点简化规则和分支搜索技术提出 $O(kn+2^k k^{2k+2})$ 的 FPT 算法.文献[28]给出了一种改进的分支搜索算法,使时间复杂性降至 $O(kn+1.324718^k k^2)$,随后,算法时间复杂性进一步降至 $O(kn+1.31951^k k^2)$ [29]和 $O(kn+1.29175^k)$ [30].在多项式空间复杂性的前提下,目前最好的算法是 Chen 等人[3]于 2001 年提出的,其时间复杂度为 $O(kn+1.285^k)$.其算法首先应用局部简化技术把图 G 中的低度点和度>k 的高度点删除,得到最多只有 k^2 条边、$2k^2$ 个点的图 G_1.对图 G_1 再应用 NT 算法(NT 算法实际上就是一个一次性的皇冠分解[31]),便可以得到大小为 $2k$ 的问题核,整个核化过程的时间复杂度为 $O(kn+k^3)$.

Clique Cover 问题在编译优化、计算几何和应用统计学等许多领域都有很重要的应用[14,32].该问题是 NP 完全问题,目前有很多求解这一问题的启发式算法[32−34].Gramm 等人[35]采用一种新的数据结构,设计了一个时间复杂度为 $O(nm)$的启发式算法.参数化 Clique Cover 问题可以表述为:给定无向图 $G=(V,E)$和正整数 k,图 G 中是否有至多 k 个团,使得 G 的每条边的两个端点至少落在这 k 个团之一呢?文献[35]证明了 k-Clique Cover 是 FPT 问题.其算法应用局部简化技术,提出了 3 条简化规则,在 $O(n^4)$时间内得到了核大小为 2^k 的新实例.

Hyperplane Cover 问题:给定 d 维实数空间里的 n 个点组成的集合 S,若能够找到 k 个超面,使得 S 中的每个点至少落在其中一个超面上,则称点集 S 能够被 k 个超面所覆盖.这个问题即便是线覆盖问题(即 $d=2$,用 k 条直线来覆盖点集 S),也是 NP 难解的[36].在具体应用中,找 k 个球面来覆盖点集或找 k 个多项式函数来覆盖点集等实际问题,都可以看作是超面覆盖的特例[37].若给定超面覆盖问题的一个实例(S,k,d),可通过降维方法将 S 中的每个点划分到 k 个子集 $S_1,S_2,\dots,S_k$ 中,然后应用边界搜索树算法进行递归搜索,最终得到问题实例(S,k,d)的解[37].这一算法的时间复杂度为 $O^*(k^{dk}n)$,从中可以看出,超面覆盖问题是一个以 d 和 k 为参数的 FPT 问题.Langerma 等人[37]提出了超面覆盖的核心化算法,该算法基于局部简化技术.通过两个简化规则,将原问题转化为核大小为 $O(k^d)$的等价问题,再应用边界搜索树技术得到时间复杂度为 $O^*(k^{d(k+1)}+n^{d+1})$的 FPT 算法.

3.2 Packing类问题

Set Packing 问题广泛应用于调度、代码优化和生物信息学等领域.文献[38]对 Set Packing 问题的多种表现形式及其研究作了较为系统的讨论.Set Packing 的一般化问题简称为 GSP 问题,即给定一个含有 n 个元组的集合 $S=(S_1,\dots,S_n)$,S_i 中的每个元素均取自符号集 U,目标是在 S 中寻找一个最大子集 S',使 S'中的任何两个元组之间均没有相同元素.在实际应用中,当我们限定 S 中每个元组至多由 m 个元素组成时,便称为 m-SP 问题;若考虑 S 中各个元组在整体中的作用不同,因此将其赋上不同大小的权值时,便称为带权 m-WSP 问题[39,40].GSP 问题可以在多项式时间内转化成最大独立集和最大团问题[5].Fomin 应用加权分治技术将最大独立集问题的时间复杂度降至 $O(1.2209^n)$[41].对于 m-SP 问题,当 $m=2$ 的时候,存在多项式时间算法[42].当 $m\geq 3$ 时是 NP 难解的,m-SP 和

© 中国科学院软件研究所 http://www.jos.org.cn

m-WSP 的近似算法主要是基于贪婪思想并加以局部改进[43,44].

参数化的 GSP 属于 W[1]完全问题,不存在 FPT 算法,除非 P=NP.参数化 m-SP 是 FPT 问题[39],m=3 时称为 3-SP 问题,这是参数计算中研究的热点问题之一.核心化、局部贪婪、动态规划、彩色编码等算法技术在 3-SP 问题上均得到了成功的应用.文献[45]提出了 3-SP 的核心化算法,用一些特殊的模式三元组与 S 中的所有三元组来匹配,如果匹配,就删掉这部分三元组.完成这一核心化过程后,可以得到大小为 $O(k^3)$的核.文献[46]提出一个时间复杂度为 $O^*(4.61^{3k})$算法,该算法首先应用彩色编码技术得到 $O^*(6.1^k)$的着色机制,再通过动态规划方法进行求解.文献[47]基于随机分治技术提出了 $O^*(3.52^{3k})$的确定性参数算法,这是目前关于此问题的最好结果.

图的 H-Packing[48]:给定两个连通图 G 和 H,图 G 的一个 H-Packing 是 G 的一个互不相交的子图的集合,每个子图都与图 H 同构.H-Packing 问题广泛应用于调度、无线传感器跟踪、线路板设计和代码优化等领域,如果图 H 是一个至少有 3 个点的连通图,这个问题便是一个 NP 完全问题了.针对 H 的不同类型,图的 H-Packing 可转化为一些特殊问题,如 H 为完全二分图、路径、三角形和圈等[14,17,34,49].

在参数理论中,当 H 为一般图时,Fellows 等人[17]给出了 H 为一般图时目前最好的研究成果,其时间复杂度为 $2^{O(|H|k\log k+k|H|\log|H|)}$.当 H 为三角形时,称为 k-边不相交三角形 packing 问题,该问题即使对于平面图来说也是 NP 完全问题[50].Fellows 等人[51]应用皇冠分解技术,对 k-边不相交三角形 packing 问题设计出核为 $O(k^3)$的 FPT 算法.Mathieson 等人[14]进一步在局部简化和皇冠分解的基础上,应用极值归纳技术进行核心化,设计出 $O(2^{4.5k\log k+4.5k})$的核心化算法,核大小为 $4k$.若 H 为完全二分图 $K_{1,s}$,称为 k-$K_{1,s}$-Packing 问题.Prieto 等人[21]应用皇冠分解技术,在 $O(n)$时间内得到 $O(k^2)$的问题核.同时,该文对于 k-$K_{1,s}$-Packing 的一个特例 k-P_2-packing 问题(判断图 G 中是否存在 k 条由 3 个点和 2 条边组成的一条路径),应用皇冠分解技术和极值归纳技术提出了 $O^*(2^{5.301k})$的核心化算法,核大小为 $15k$.2008 年,Wang 等人[16,52]在文献[21]的核心化算法基础上提出新的简化规则,使问题核降为 $7k$,算法的时间复杂度也降到 $O^*(2^{4.142k})$,是目前对 k-P_2-packing 问题的核化效果最好的算法.

3.3 Cut类问题

k-Max Cut 问题:图 G 中的顶点集 V 能否划分为两个子集 V'和 V'',使得从 V'到 V''的边数不少于 k 条?Max Cut 在电路布局设计和统计物理等方面应用广泛[53],该问题是 NP 完全的,其近似算法取得了很多成果[54].1999 年,Mahajan 给出了 k-Max Cut 问题的第 1 个 FPT 算法[55],算法时间复杂度为 $O(|E|+|V|+4^k k)$.随后,Fedin 等人[56]应用局部简化技术和 splitting 转化规则,给出了 $O(\text{poly}|E|)\times 2^{|E|/4}$的核心化算法,核为 $O(2^{|E|/4})$.目前,该问题最好的 FPT 算法是 Prieto 所提出的核心化算法[53].算法提出 5 个简化规则,使实例在 $O^*(n^2)$时间内得到简化,再应用极值归纳技术,得到至多 k 个顶点、$2k$ 条边的核,整个算法的时间复杂度为 $O(k\times 2^{k/2}+n^2)$.

Multicut 问题:图的多割问题是计算机和通信网络可靠性和鲁棒性研究方面的基本模型,此外,Multiway cut 是多割的一个特例问题,而 Multiflow 是多割的对偶问题,因此,Multicut 是一个很重要的应用问题.对于给定的图 $G(V,E)$和端点的点对集 $H\subseteq V\times V$,找到一个最小的边集或点集 S,使得将 S 删除后,H 中的点对不存在路径.这样的 S 集便是多割的解.若以 H 中的顶点对数 k 为参数,则可转化为以 k 为参数的 k-Multicut 问题.

根据图 G 是有向图、无向图还是树,Multicut 可进一步细分为 Di-Multicut、Undi-Multicut 和 Tree-Multicut.对于点对集 H 中的点是否可作为解的一部分选入 S 集,又可把问题划分为不受限的 Multicut 问题(UVMC)和受限的 Multicut 问题(RVMC).根据选入 S 集中的元素是边还是点,Multicut 问题又划分为边-Multicut 和点-Multicut 问题.对 Multicut 的各种各样的变化问题的难度,Costa 等人在文献[57]中加以了总结.对于无向 Multicut,若 k=2,是多项式可解的,若 $k\geq 3$,是 Max SNP-hard 的,即不存在 PTAS 算法;对于有向 Multicut 问题,若 k=2,是 NP-hard 的,若 $k\geq 3$,是 Max SNP-hard 的,该问题目前最好的近似算法是 $O(\min\{\text{opt}, \sqrt{n}\})$[58].对于一般图,UVMC 和 RVMC 均是 NP 难的.若将图限制为树,则 UVMC 是 P 类问题而 RVMC 是 NP 难的[59].从参数的角度来看,Multicut 可以有 3 个参数加以选择或组合:至多允许删除的边的条数或点的个数 k、H 的大小 l 和图 G 的树宽ω.若仅仅考虑参数 k,则对于一般图的 UVMC 和 RVMC 问题是否是 FPT 的,目前还是开放性问题.若将问题限制为树,则 UVMC 是 P 类问题,而 RVMC 是 NP 完全的.对于一般图的 UVMC 问题,Marx[60]应用图分割技术,给出一个

$O(2^{kl}k^k4^{k^3}|G|^{O(1)})$ 的算法,证明了一般图的 UVMC 问题是双参数 FPT 的.对于树的 RVMC 问题,Guo 等人[61]结合边界搜索树算法,设计出 $O(2^k\cdot|V|\cdot l)$ 的双参数 FPT 算法.此外,Guo 等人[62]针对树的 RVMC 问题,应用局部简化技术提出 8 个简化规则,可在 $O(2^kmn)$时间内得到大小为 $O(k^{3k})$的核.从核大小来看,这一结果并不比文献[61]所采用的搜索树算法更高效,但作为树的 Multicut 问题的第 1 个核心化算法,为我们提供了问题核的上界.

3.4 其他参数化问题

核心化技术已成功地应用于求解许多重要的参数化问题.针对一些典型的 FPT 问题,表 1 总结了在这类问题的 FPT 算法设计过程中所采用的核心化技术、核化的大小以及核心化算法的时间复杂性.通过表 1 可以看出,在对问题进行核心化的阶段,局部简化技术是应用最普遍、也是最基本的,几乎每一个参数化问题都会根据问题的不同特点,设计出局部简化规则.此外,多种核心化技术往往会结合在一起,以求得到问题的更小些的核.

Table 1 Typical FPT problems: Kernelization technique used and its kernel size

表 1 典型的 FPT 问题核心化技术和核大小

FPT questions	References	Kernelization techniques					Time complexity of Kernelization algorithms
		Crown reduction	Extremal induction	Randomized method	Local reduction	Kernel size	
Vetex cover	Faisal[15]	√			√	$3k$	$O(n^{5/2})$
	Chen[3]	√			√	$2k$	$O(kn+k^3)$
3-Hitting set	Niedermeier[63]				√	$O(k^3)$	$O(n)$
	Nishimura[64]				√	$6k^2$	$O(kn+k^4)$
Unweighted set splitting	Dehne[65]		√		√	$2k$	$O(n^4)$
	Dehne[7]	√			√	k–1	Poly(n,k)
	Chen[20]			√	√	$2k$	$O(n+2k^2)$
Weighted set splitting	Chen[22]			√	√	$2k$	$O(n+2k^2)$
k-Maximum cut	Prieto [53]		√		√	k	$O(n^2)$
Planar dominating set	Alber[66]				√	$335k$	$O(n^3)$
	Chen[67]				√	$67k$	$O(n^3)$
edge disjoint triangles packing	Mathieson[14]	√	√		√	$4k$	$O(n)$
Vertex disjoint triangles packing	Fellows[17]	√			√	$O(k^3)$	$O(nk+n^2)$
$K_{1,s}$-Packing	Prieto[21]	√	√		√	$O(k^2)$	$O(n)$
P_2-Packing	Wang [16,52]	√			√	$7k$	$O(nk)$
Cluster editing	Guo[19]		√		√	$4k$	$O(nm^2)$
k-Internal spanning tree	Prieto[20]		√		√	$O(k^3)$	$O(nk^2)$
Clique cover	Gramm[35]				√	2^k	$O(n^4)$
Dim-Set-Cover	Langerman[37]				√	$O(k^d)$	$O(n^{d+1})$
MaxSat	Niedermeier[68]				√	$O(k^2)$	$O(n)$
Clique partition	Mujuni[69]				√	k^2	$O(n^2)$
Connectivity augmentation	Guo[70]				√	$O(k^2)$	$O(mn+n^2)$
Undirected MaxLeaf	Bonsma[71]				√	$4k$	Poly(n,k)
NonBlocker	Dehne[72]				√	$5k/3$	Poly(n,k)

4 结论与进一步的研究

核心化技术是参数算法设计领域中一种最基本的算法设计技术,一直以来是参数计算领域的研究热点之一.针对 FPT 问题,参数理论有两个追求的目标:一是想方设法设计出时间效率更高的算法;二是想方设法找出更小的问题核.对这两个努力目标来说,核心化技术都是最为基本的.本文分别阐述了 4 种主要的核心化技术的算法思想和在实际应用中的基本核化过程.通过对这些算法思想的分析和比较,使人们对核心化技术有一个更全面的理解,同时也可为其他参数问题的核心化算法提供求解的思路.核心化技术在理论和应用等领域依然存在许多有待我们进一步研究的问题和挑战,主要体现在以下几个方面:

(1) FPT 问题核的下界研究

FPT 问题核的下界研究是参数理论在最近几年才开始的一个新的研究方向,核的下界研究无论在理论方

© 中国科学院软件研究所 http://www.jos.org.cn

面还是应用方面都有实际的指导意义.我们总是希望 FPT 问题经过核心化处理后可以得到多项式核甚至线性核,但是,并不是所有的 FPT 问题均存在多项式核.Chen 等人在文献[67]中针对可线性核化的平面图问题设计了对偶参数方法,从而证明了许多平面图问题的核下界.最近,Bodlaender 等人[73]进一步证明了许多的 FPT 问题(如 k-path,k-cycle,k-cutwidth,k-treewidth 等问题)不存在多项式核.

FPT 问题的下界在理论和应用上都存在许多有待进一步研究的问题.从应用角度看,主要体现在以下几个方面:① 亚指数级的核下界证明问题.对于已经证明不存在多项式核的 FPT 问题,能否证明其存在亚指数级的核下界?② 多项式核下界存在性证明问题.有许多应用广泛的问题,如有向图反馈集、Clique Cover、Graph Bipartition、有向最大叶等问题,是否存在多项式核下界?③ 多项式核下界不存在性证明问题.例如,有许多 NP 完全的图问题往往适于以树宽为参数,却至今未能找到多项式核的核心化算法,是否以树宽为参数的图问题不存在多项式核下界?

(2) FPT 问题证明及其核优化的研究

目前,依然存在许多的问题尚未能证明其是否是固定参数可解的,如多项式识别测试问题、以树宽为参数的图同构问题等[74].对于未知其参数复杂性的问题,若我们能够应用核心化技术找到其问题核,则说明该问题是固定参数可解的了.此外,针对 FPT 问题,进一步设计出问题核更小的算法,是参数理论的努力目标.例如,3-set packing 问题和无向图反馈集问题已有 $O(k^3)$的核心化算法,是否能够找到比 $O(k^3)$更小的核呢?基于我们的研究基础,对于 3-set packing 问题,我们认为可以尝试从另一个角度对该问题的核进行优化.观察可知:k^3 的集合可由最少 $3k$ 个元素构成;如果 k^3 个集合含有 $3k^3$ 个元素,则问题可在多项式时间内被求解.基于上述两个观察,我们将考虑问题实例中元素个数的界,即从元素的角度对问题的大小进行限制,进而设计出 3-set packing 更小的核的核心化算法.而对于 P_2-packing 问题,目前最小的核为 $7k$.在对该问题核心化的过程中,文献[16]中只考虑了 Q_0 点与 Q_1 点的简单连接关系,且以降低 Q_0 点的个数为目标.我们能否基于局部贪婪技术,并结合皇冠分解技术,进一步探讨 P_2-packing augmentation 问题的核?亦即给定一个 k-P_2-packing,如果图中存在$(k+1)$-P_2-packing,则该图的大小最多为多少?

(3) FPT 核心化新技术的研究

随着研究的越来越深入,FPT 的核心化新技术也不断地涌现,尤其是传统的算法设计技术日益融合到核心化技术中来.随机方法便成功地运用于求 FPT 问题的核.有许多的生物信息学问题如生物分子相互作用网络问题,其元素具有不同程度的随机性和不确定性,研究运用随机方法对这类问题加以核化,进而得到拓扑结构中心域的元素核,将有助于关键蛋白质的识别和分析.

从归纳技术的角度来看,目前有两种很有效的算法技术,即递归压缩技术和极值归纳技术.递归压缩技术适于求解具有参数单调性的最小化参数问题,即求小于等于 k 的参数问题.其核心思想是将(G,k)实例转化为$(G-v,k)$实例,使问题 G 的规模逐步降低.极值归纳技术适于求解具有参数单调性的最大化参数问题,即求大于等于 k 的参数问题.其核心思想是从$(G,k-1)$实例出发,逐步归纳出(G,k)实例的解.但是,相比于$(G-v,k)$和$(G,k-1)$,还有一个更高效的实例即$(G-v,k-1)$,目前还没有相应的递归或归纳思想的算法设计技术,有待进一步的分析研究.

(4) 核心化技术在复杂网络等领域的应用研究

核化技术目前主要应用于确定性问题的研究,从图论的角度来说,一般都是研究具有明确拓扑结构的图.许多实际问题,如生物分子相互作用网在不同的细胞或同一细胞的不同生活时刻是不完全相同的,而互联网的拓扑结构时刻也在动态变化之中.这些问题都具有不同程度的随机性和不确定性,一般都表现出复杂网络的特性.复杂网络的性能指标与其拓扑结构密切相关,拓扑结构演化时,其性能指标会受到一定的影响.目前,许多拓扑结构优化的目标是追求性能指标对拓扑结构具有的相对不敏感性(稳定性);复杂网络的动态、策略控制是阶段性地改变拓扑结构和系统的关键参数,使得整个网络的目标函数达到最优或次优.基于这样的考虑,通过核化,将核引申为网络(图)中对某些操作或过程具有相对稳定性的结构,同时也可以反映网络拓扑随着其中具有某种特性的点或局部结构的缺失而波动的范围以及程度等方面的情况,从而对复杂网络的动态及其系统动力学演变规律进行研究.

© 中国科学院软件研究所 http://www.jos.org.cn

(5) 核心化技术在非参数理论领域的应用研究

核心化思想的本质就是对问题加以预处理,这一思想在非参数理论领域也得到了广泛地应用.在求解非参数化问题的过程中,往往针对问题的性质特征,对数据加以去伪存真,过滤干扰数据或次要数据,从而使得问题的规模得到极大的简化.例如,在机器学习领域,数据集往往是高维空间数据,其基本预处理过程是通过分析数据本身的结构特征,并针对结构特征进行选择,以达到降维的目的.如果对数据不进行任何的预处理,仅仅通过特征抽取算法找出特征集,则由于高维等原因,反而会导致问题的精确度下降.但是,目前非参数理论领域的预处理算法主要还是依赖于贪婪选择,无法保证数据的质量.在实际应用中,我们针对关键蛋白质网络问题,应用参数理论的局部简化技术去除非关键结点,使问题规模得到了极大地降低.我们认为,把参数理论的核心化技术应用于非参数理论领域的问题的预处理过程,是值得研究的一个方向.

References:

[1] Chen J. Parameterized computation and complexity: A new approach dealing with NP-hardness. Journal of Computer Science and Technology, 2005,20(1):18−37.

[2] Downey RG, Fellows MR. Parameterized computational feasibility. Feasible Mathematics II, Birkhauser, 1992. http://www.mcs.vuw.ac.nz/math/papers/feasible.ps

[3] Chen J, Kanj IA, Jia W. Vertex cover: Further observations and further improvements. Journal of Algorithms, 2001,41:280−301.

[4] Roth-Korostensky C. Algorithm for building multiple sequence alignments and evolutionary trees [Ph.D. Thesis]. Diss.ETH No.13550, 2000.

[5] Garey MR, Johnson DS. Computers and Intractability: A Guide to the Theory of NP-Completeness. New York: W.H.Freeman & Co., 1979.

[6] Chen J, Friesen DK, Jia WJ, Kanj IA. Using nondeterminism to design efficient deterministic algorithms. Algorithmica 40, 2004. 83−97.

[7] Dehne F, Fellows M, Rosamond F, Shaw P. Greedy localization, iterative compression and modeled crown reductions: New FPT techniques and improved algorithms for max set splitting, and a novel 2*k* kernelization for vertex cover. LNCS 3162, 2004. 271−281.

[8] Jia WJ, Zhang CL, Chen J. An efficient parameterized algorithm for m-set packing. Journal of Algorithms, 2004,50(1):106−117.

[9] Alon N, Yuster R, Zwick U. Color-Coding. Journal of the ACM, 1995,42(4):844−856.

[10] Sloper C, Telle JA. An overview of techniques for designing parameterized algorithms. The Computer Journal, 2008,51:122−136.

[11] Wang JX, Huang YN, Chen JE. A motif finding algorithm based on color coding technology. Journal of Software, 2007,18(6): 1298−1307 (in Chinese with English abstract). http://www.jos.org.cn/1000- 9825/18/1298.htm

[12] Wang JX, Liu YL, Chen JE. An effective coloring algorithm for close relationship between the scales of element set and color set and its application. Chinese Journal of Computers, 2008,31(1):32−42 (in Chinese with English abstract).

[13] Fellows M. Blow-ups, win/win's, and crown rules: some new directions in FPT. LCNS 2880, Berlin: Springer-Verlag, 2003. 1−12

[14] Mathieson L, Prieto E, Shaw P. Packing edge disjoint triangles: A parameterized view. LNCS 3162, 2004. 127−137.

[15] Abu-Khzam FN, Collins RL, Fellows MR, Langston MA, Suters WH, Symons CT. Kernelization algorithms for the vertex cover problem: Theory and experiments. ALENEX/ANALC, 2004.62−69.

[16] Wang JX, Ning D, Feng QL, Chen JE. An improved parameterized algorithm for a generalized matching problem. In: Agrawal M, *et al*., eds. Proc. of the 5th Annual Conf. on Theory and Applications of Models of Computation (TAMC2008). LNCS 4978, Berlin: Springer-Verlag, 2008. 212−222.

[17] Fellows M, Heggernes P, Rosamond F, Sloper C, Telle JA. Finding *k* disjoint triangles in an arbitrary graph. LNCS 3353, 2004. 235−244.

[18] Fellows MR, McCartin C, Rosamond FA, Stege U. Coordinatized kernels and catalytic reductions: An improved FPT algorithm for max leaf spanning tree and other problems. LNCS 1974, 2000. 240−251.

[19] Guo J. A more effective linear kernelization for cluster editting. LNCS 4614, 2007. 36−47.

© 中国科学院软件研究所 http://www.jos.org.cn

[20] Prieto E, Sloper C. Reducing to independent set structure-the case of *k*-internal spanning tree. Nordic Journal of Computing, 2005,12(3):308−318.

[21] Prieto E, Sloper C. Looking at the stars. Theoretical Computer Science 351, 2006. 437−445.

[22] Chen J, Lu S. Improved parameterized set splitting algorithms: A probabilistic approach. Algorithmica, 2008.

[23] Zhang HJ, Ling CX. An improved learning algorithm for augmented naive bayes. LNCS 2035, 2001. 581−586.

[24] Dehne F, Fellows MR, Rosamond FA. An FPT algorithm for set splitting. LNCS 2880, 2003. 180−191.

[25] Lokshtanov D, Sloper C. Fixed parameter set splitting, linear kernel and improved running time. In: Proc. of the Algorithms and Complexity in Durham 2005. King's College Press, 2005.105−113.

[26] Chang L. The parameterized vertex problem and the minimum vertex cover problem [MS. Thesis]. Changsha: Central South University, 2007 (in Chinese with English abstract).

[27] Buss JF, Goldsmith J. Nondeterminism within P. SIAM Journal on Computing, 1993,22(4):560−572.

[28] Stege U, Fellows M. An improved fixed parameter algorithm for vertex cover. Technical Report, ETH Zurich: Department of Computer Science, 1999.

[29] Downey RG, Fellows MR, Stege U. Parameterized complexity: A framework for systematically confronting computational intractability. In: Abello J, Vitter J, eds. Proc. of the AMS(DIMACS). 1999. 49−99.

[30] Niedermeier R, Rossmanith P. Upper bounds for vertex cover further improved. LNCS 1563, 1999. 561−570.

[31] Chlebfk M, Chlebikova J. Crown reductions for the minimum weighted vertex cover problem. Electrunic Colloquium on Computational Complexity, 2004.

[32] Rajagopalan S, Vachharajani M, Malik S. Handling irregular ILP within conventional VLIW chedulers using artificial resource constraints. In: Chair T, *et. al*., eds. Proc. of the CASES. ACM Press, 2000. 157−164.

[33] Piepho HP. An algorithm for a letter-based representation of all-pairwise comparisons. Journal of Computational and Graphical Statistics, 2004,13(2):456−466.

[34] CaPrara A, Paneonesi A, Rizzi R. Packing cycles in undirected graphs. Algorithms, 2003,48(l):239−256.

[35] Gramm J, Guo J, Huffner F, Niedermeier R. Data reduction, exact, and heuristic algorithms for clique cover. In: Raman R, ed. Proc. of the 8th ACM-SIAM ALENEX. ACM-SIAM, 2006. 86−94.

[36] Megiddo N, Tamir A. On the complexity of locating linear facilities in the plane. Operations Research Letters, 1982,1:194−197, 662, 671.

[37] Langerman S, Morin P. Covering things with things. LNCS 2461, 2002. 662−674.

[38] Ma ZY. Research and application of measure and conquer in set packing [MS. Thesis]. Changsha: Central South University, 2007 (in Chinese with English abstract).

[39] Cesati M. Compendium of parameterized problems. 2001. http://citeseer.ist.psu.edu/ viewdoc/cesati01compendium.html

[40] Arkin EM, Hassin R. On local search for weighted k-set packing. LNCS 1284, 1997. 13−22.

[41] Fomin FV, Grandoni F, Kratsch D. Measure and conquer: A simple $O(2^{0.288n})$ independent set algorithm. In: Stein C, ed. Proc. of the 17th Annual ACM-SIAM Symp. on Discrete Algorithm. 2006. 18−25.

[42] Gabow H. An efficient implementation of Edmonds' algorithm for maximum matching on graphs. STAN-CS-72-328, Technical Report, No.31, Stanford University, 1972.

[43] Hazan E, Safra S, Schwartz O. On the Complexity of Approximating *k*-Set Packing. Computational Complexity, 2006. 20−39.

[44] Berman P. A $d/2$ approximation for maximum weight independent set in d-claw free graphs. LNCS 1851, 2000. 214−219.

[45] Fellows MR, Knauer C, Nishimura N, Ragde P. Faster fixed-parameter tractable algorithms for matching and packing problems. LN CS 3221, 2004. 311−322.

[46] Liu Y, Lu S, CJ, Sze SH. Greedy localization and color-coding:improved matching and packing algorithms. LNCS 4169, 2006. 84−95.

[47] Wang JX, Feng QL. An $O^*(3.52^{3k})$ parameterized algorithm for 3-set packing. In: Agrawa M, *et al*., eds. Proc. of the Theory and Applications of Models of Computation (TAMC 2008). LNCS 4978, 2008. 82−93.

[48] Ning D. Parameterized algorithm for the matching and packing problem [MS. Thesis]. Changsha: Central South University, 2007 (in Chinese with English abstract).

© 中国科学院软件研究所 http://www.jos.org.cn

[49] Hassin R, Rublnstein S. An approximation algorithm for maximum packing of 3-edge Paths. Information processing Letters, 1997, 63−67.

[50] Holyer I. The NP-completeness of some edge-partition problems. SIAM Journal of Computer, 1981, 713−717.

[51] Fellows M, Heggernes P, Rosamond F, Sloper C, Telle JA. Finding k disjoint triangles in an arbitrary graph. In: Hromkovic J, *et al*., eds. Proc. of the WG 2004. 2004. 235−244.

[52] Wang JX, Ning D, Feng QL, Chen JE. An improved parameterized algorithm for P_2-packing problem. Journal of Software, 2008,19(11):2879−2886 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/19/2879.htm

[53] Prieto E. The method of extremal structure on the k-MAXIMUM CUT Problem. Proc. of Computing: Australian Computer Science Communications, 2005,27(4):119−126.

[54] Paradimitriou CH, Yannakakis M. Optimization, approximation, and complexity classes. Journal of Computer and System Sciences 43, 1991. 425−440.

[55] Mahajan M, Raman V. Parameterizing above guaranteed values: MaxSat and MaxCut. Journal of Algorithms, 1999,31(2):335−354.

[56] Fedin SS, Kulikov AS. A $2^{|E|/4}$-time algorithm for MAX-CUT. Zapiski Nauchnyh Seminarov POMI, 2002,293:129−138.

[57] Costa MC, Létocart L, Roupin F. Minimal multicut and maximal integer multiflow: A survey. European Journal of Operation Research, 2005,162:55−69.

[58] Gupta A. Improved results for directed multicut. In: Farach-Colton M, ed. Proc. of the 14th annual ACM-SIAM Symp. on Discrete Algorithms (SODA 2003). Society for Industrial & Applied Mathematics, 2003. 454−455.

[59] Calinescu G, Fernandes CG, Reed B. Multicuts in unweighted graphs and digraphs with bounded degree and bounded tree-width. Journal of Algorithms, 2003,48:333−359.

[60] Marx D. Parameterized graph separation problems. In: Downey R, *et al*., eds. Proc. of the 1st IWPEC. LNCS 3162, Springer-Verlag, 2004. 71−82.

[61] Guo J, Hüffner F, Kenar E, Niedermeier R, Uhlmann J. Complexity and exact algorithms for multicut. LNCS 3831, Springer-Verlag, 2006. 303−312.

[62] Guo J. Rolf Niedermeier. Fixed-Parameter tractability and data reduction for multicut in trees. Networks, 2005,46(3):124−135.

[63] Niedermeier R, Rossmanith P. An efficient fixed-parameter algorithm for 3-hitting set. Journal of Discrete Algorithms, 2003,1:89−102.

[64] Nishimura N, Ragde P, Thilikos DM. Smaller kernels for hitting set problems of constant arity. In: Downey R, *et al*., eds. Proc. of the IWPEC 2004. Springer-Verlag, 2004. 121−126.

[65] Dehne F, Fellows MR, Rosamond FA. An FPT algorithm for set splitting. LNCS 2880, 2003. 180−191.

[66] Alber J, Fellows M, Niedermeier R. Polynomial time data reduction for dominating set. Joural of the ACM 51, 2004. 363−384.

[67] Chen J, Fernau H, Kanj IA, Xia G. Parametric duality and kernelization: Lower bounds and upper bounds on kernel size. STACS 2005. 269−280.

[68] Niedermeier R. Ubiquitous parameterized-invitation to fixed parameter algorithms. LNCS 3153, Springer-Verlag, 2004. 84−103.

[69] Mujuni E, Rosamond F. Parameterized complexity of the clique partition problem. In: Harland J, *et al*., eds. Proc. of the CATS 2008. 2008.

[70] Guo J, Uhlmann J. Kernelization and complexity results for connectivity augmentation problems. In: Dehne F, *et al*., eds. Proc. of the WADS 2007. LNCS 4619, 2007. 484−495.

[71] Bonsma P, Zickfeld F. Spanning trees with many leaves in graphs without diamonds and blossoms. In: Eduardo SL, *et al*., eds. Proc. of the LATIN 2008. 2008. 531−543.

[72] Dehne F, Fellows M, Fernau H, Prieto E, Rosamond F. Nonblocker: Parameterized algorithms for minimum dominating set. In: Wiedermann J, *et al*., eds. Proc. of the SOFSEM 2006. 2006.

[73] Bodlaender HL, Downey RG, Fellows MR, Hermelin D. On problems without polynomial kernels. In: Proc. of the ICALP(1)2008. 2008. 563−574. http://www.informatik.uni-trier.de/~ley/db/conf/icalp/icalp2008-1.html#BodlaenderDFH08

[74] Bodlaender HL, Demaine ED, Fellows MR, Guo J, Hermelin D, Lokshtanov D, Muller M, Raman V, van Rooij J, Rosamond FA. Open problems in parameterized and exact computation—IWPEC 2008. Technical Report, CS-2008-017, 2008.

© 中国科学院软件研究所 http://www.jos.org.cn

附中文参考文献:

[11] 王建新,黄元南,陈建二.一种基于彩色编码技术的基序发现算法.软件学报,2007,18(6):1298−1307. http://www.jos.org.cn/1000-9825/18/1298.htm

[12] 王建新,刘云龙.陈建二.一种在元素与颜色规模相近时的有效着色算法及其应用.计算机学报,2008,31(1):32−42.

[26] 常乐.参数化点覆盖及最小点覆盖问题研究[硕士学位论文].长沙:中南大学,2007.

[38] 马振宇.加权分治技术在 Set Packing 问题中的应用与研究[硕士学位论文].长沙:中南大学,2007.

[48] 宁丹.Matching 和 Packing 问题的参数算法研究[硕士学位论文].长沙:中南大学,2007.

[52] 王建新,宁丹,冯启龙,陈建二.P2-Packing 问题参数算法的改进.软件学报,2008,19(11):2879−2886. http://www.jos.org.cn/1000-9825/19/2879.htm

李绍华(1964－),男,江西南康人,博士生,副教授,主要研究领域为参数计算理论与算法.

冯启龙(1982－),男,博士生,主要研究领域为参数计算理论与算法.

王建新(1969－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机优化算法,网络优化理论,生物信息学.

陈建二(1954－),男,博士,教授,博士生导师,主要研究领域为生物信息学,计算机理论,计算复杂性及优化,计算机网络优化算法,计算机图形理论与算法.

© 中国科学院软件研究所 http://www.jos.org.cn