**通过局部改进逼近离散集合**

Magntús M. Halldórsson

**摘要：**我们考虑了一种古老的优化技术，并将其应用于离散项集合的近似。这种技术属于局部改进搜索: 尝试通过添加一些项而删除另一些项来扩展集合。

通过NC算法或在近似线性序列时间内，得到了在k+1无爪图中，k-DM、k-SET PACKING和独立集的(k + 1)/2性能比。它可以推广到k/2+ε(当 k≥4) ，或推广到(k+1)/3+ε(当最大度有界时) 。

我们还得到了诱导子图问题、有界度图中的独立集、无爪图中的顶点覆盖、集覆盖和互补目标函数下的图着色等问题的的改进性能比。

**1 前言**

局部搜索是解决较困难的组合优化问题的常用工具。它有多种方法，包括贪心搜索、爬山、模拟退火、随机爬山、增广路和局部变化。然而，尽管在实际情况中取得了相当可观的成功，但几乎没有算法在最坏情况下表现积极。

普遍认为，局部搜索在理论上很大程度是无用的。在形式上这被称作PLS完全（PLS-completence）[14]，换言之:“最难的局部优化问题”。PCS完全问题及其近似问题们（邻域？）覆盖了大多数流行的局部搜索：关于旅行商问题的*t*-opt算法和Lin-Kernighan算法；关于图分割问题Kernighan-Lin算法和Fiduccia-Matheyses算法；关于最大割问题的FLIP算法。PLS完全问题上的局部搜索算法能否在最坏情况下的多项式时间收敛于局部最优解。多种有指数时间复杂度上限的常用启发式算法已经被建立。在一些案例中，强得多的复杂度结果被展现：对给定实例和起始解，没有任一序列的优化以小于指数长时间的局部最优解结束[28]。这是有权重问题的例子；对无权重的PLS完全问题，获取局部最优解是P-完全问题。[26][14][28]

最坏情况下的局部搜索真正的问题是，几种在相对简单的问题上都有很好平均表现的流行的启发式算法，都被证明存在弱点。例如，Metropolis算法，模拟退火的核心算法，不仅不能找到极大团[12](这是困难问题) ，也在最大匹配问题上表现糟糕[25](这不是困难问题)。对于最流行的启发式算法测试问题，度量TSP，有实例表明，流行的 t-opt 启发式算法的复杂度是Ω(n1/(2t)) [3]，比最优解糟糕。

本文的目的是在提出一些局部最优解的积极结果。目标有两个：一是提高各种优化问题的最佳性能比；二是阐明局部改进启发式算法作为近似算法的有效性。我们考虑一系列问题，它们的共同主题是它们涉及到离散项的集合，其目标是最小化或最大化集合的大小（通常定义为项的数量）。所使用的方法是自然的: 从最大/最小集合开始，然后扩大/缩小集合，直到找不到更好的只需修改少量项的解决方案为止。

我们提出以下优化性能比:

1. k/2+ε，在k+1无爪图中(k≥4)；k=3时这个数值为5/3。它通过一个NC算法实现。这个结果同样适用于 k-SET PACKING问题，k维匹配问题，H匹配问题（k=|V(H)|），以及 k-CLIQUE PACKING。
2. (*k*+1)/2，当目标是对k-SET PACKING 问题最大化集合中所有子集大小的总和时。
3. (*k*+1)/3+ε，对上一问题，当最大度数是常数时。
4. (Δ+2+1/3)/4+ε，对顶点独立集问题，当图的最大度为Δ时。
5. 同上，对于遗传诱导子集问题（HEREDITARY INDUCED SUBGRAPH），并对 INDUCED X-COLORABLE SUBGRAPH 和 VERTEX ARBORICITY 问题的性能比进行了改进。
6. 2-2fs(*k*)对于在 *k*+l无爪图中的顶点覆盖集问题，其中fs(*k*)=(*k*ln *k*-*k*+1/(*k*-1)2=2-(log k)/k(1+*o*(1))。当k=4,5时，性能比为 1.5。
7. 11/7≈1.57，对 3-集合覆盖问题（限制集合数量不超过 3）。对于 k-顶点覆盖问题，其性能比为 Hk-11/42。
8. 1.4，当目标是对图着色问题最大化|V|-c，其中c是使用的颜色数量。

在所有情况下，我们分析的启发式算法的性能上限都是紧的。这些问题都是最大SNP困难(MAX SNP-hard)的 [23][1]，因此无法在ε>0的情况下以1+ε完成近似（除非 P=NP）。对于独立集问题和其他相关问题，近似难度会随着最大度或爪数的增加而以多项式增长[6]。

**1. 1 相关工作**

我们研究的局部优化技术并不新颖。它在一个多项式时间可解的问题——最大匹配问题中可能最为常见。在这种情况下，它简化了寻找最小增广路的工作。

这里提出的分析结果中的许多想法来自于 Berman 和 Fürer[2]的工作，他们在有界限度图（bounded-degree graphs）的独立集问题极大改进了比率。Halldórsson和Radhakrishnan[9]给出了这个主题的一些变体。

**2**  离散集合和局部改进

**2. 2 局部改进**

考虑以下问题：最大化满足一定特性的离散集合，也就说它是可行的。一个解是最大的，当在当前解之外添加任何项都会破坏其可行性。然后，一个非最大解可以通过仅仅添加一个项来扩展。

设可行解为*A*。假设存在集合x,y,z，满足x∈A, y,z∉A，那么A′=(A-{x}∪{y,z})也是一个解。那么，我们称A′为A的扩展，{x,y,z}为A的2-改进。更一般地，A的改进是满足A⊕I(A和I的对称联合)是一个解，且|A⊕I|>|A|的集合I。t-改进表示增加s个新的项并去除至多s-1个项，s≤t。换言之，改进增加的项目多于删除的项目。如果一个解没有t-改进，则称其为t-局部最优解。

局部搜索从初始解开始，尝试一连串的改进，直到找不到小规模的改进。下面的架构对此进行了总结

t-opt(S)

A ← 最大解(S)

repeat

I ← A的s-改进，s ≤ t

A ← A ⊕ I

直到没有更多的改进存在

return A

**3**  装箱问题

对于装箱问题，我们称两个项冲突，当两个集合相交(在集合问题中)或者两个顶点相邻(在图问题中)。我们称一个集合是可行的，当它的项彼此没有冲突。这完全指定了这些问题的 t-OPT。

令

**定理3.1** 在 k+1无爪图和k-Set Packing中可以在实现性能比为k/2+vs的(2s)-局部最优解。当k=3, s=2时，比率为5/3。

**引理3.1** 上述问题可以在的时间内近似到k/2+ϵ，当k≥4时。

对有界度图，我们可以通过[2]的结果得到更强的性能比，尽管在时间复杂度上有显著增加。

**引理3.2** 上述问题可以在的时间内近似到(k+2)/3的因子

**更弱的研究。**我们得到的主要结果，引理3.1，通过局部改进的研究比起t-opt较不通用。我们的分析是通过限定最优解中只与我们(局部最优)解中的1或2个项相冲突的项的数量来进行的。因此，我们只需要考虑那些与当前解最多有2个项冲突的项。我们可以用图的形式来表示。

令A为当前解。我们构建一个多重图H(A)，其顶点与A中的项相对应。A之外的一个项与A 中的恰好两个项相冲突，表现在H(A)中有相应的边。如果一个项与 A 中的恰好一个项发生冲突，则意味着H(A)上的相应顶点上存在自环。

一种方式的A的改进，包含H(A)的一个诱导子图，该子图的边数比顶点数多，这样对应于这些边的项彼此之间也不会冲突。为了定理3.1的目的，我们只需要寻找一种高度限制的改进类型。它对应于一个子图，我们称之为t-ear，它是一个(非简单的)长度为 t 的路径，它在最多的 t-1顶点上，且起始边是一个自环。如果相应的项目是非冲突的，我们就得到一个t-ear-改进。

**复杂度。**迭代次数一般以 n 为界，即任何解的最大大小。主要问题是寻找改进的复杂性。

让我们首先评估寻找2-ear改进的复杂性。考虑与 A 中的恰好一个顶点相冲突的项的集合。与 A 中的同一个顶点相冲突的任何一对非冲突项都构成一个改进。我们检查每对项的时间复杂度时O(n2)；如果给定冲突图，则最多检查每条边一次，时间复杂度为O(n+m)。

对于问题的集合版本，可以更快地找到改进。对于3-SET PACKING的2-opt 情形，考虑仅与一个相同的A相邻的集合，抛弃每个集合的关联元素，得到一个2-集集合，其中我们寻找一个不相交对。我们遍历整个集合，将每个集合与最多包含前三个集合的短列表进行比较。该列表包含一个星形(最多三个集合，共有一个项目)或一个三角形(三个元素集合中的所有三对) ; 任何附加集合要么与存储集合中的一个不相交，要么包含星形的固定元素。因此，这个过程可以通过在 O (n)时间内对列表进行简单扫描来完成。因此，一个2-opt的迭代可以在输入大小的线性时间内被找到。

为了寻找一个t-ear改进，我们在H中检查了每条长度为 t-2的简单路径，其中可能在任一端存在自环。路径的数量受限于H中t距离邻居的数量，或者说Δt-1。因此改进搜索的时间复杂度上限是O(mΔt-2)。对set packing问题，它可以被改进到O(nkkΔt-2)

**3. 4 有界度图中的顶点独立集**

顶点独立集即使在最大度为3时仍保持NP完全性。独立集问题的有界度情形，可以看作是无爪情形的一个约束。然而，它也允许一些额外的情形，这样性能比会大大提高。Berman和Fürer[2]介绍了一种有(Δ+3)/5+ε性能比的局部改进方法，但有着爆炸的时间复杂度。在寻找可行方法的研究中，一个性能比更弱的研究题出可以在含Δ常量的线性时间里实现(Δ+3)/4的比率[9]。我们进一步的研究将其提升到(Δ+2)/4+ε(Δ为偶数)或(Δ+2+1/3)/4+ε（Δ为奇数）。

算法的改进在于：在当前解中将它递归地应用与至少两个顶点相邻的顶点所诱导的图。这个子图的最大度等于2，会使近似变得容易一些。我们保留它被证明大于我们旧解的结果，并重复它和之前的局部改进过程，直到解在下面两个标准下是最佳的：

(3.10) b0 ≤ ρΔ-2a

当ρd是最大度为d的图的性能比。

我们还可以得到(3.8)的变体，在2sΔ邻域下：

(3.11) b0 ≤ (Δ+vs)a0/2

因为a0的t-改进在b0中有t个顶点，它们最多与A中的tΔ个顶点相邻。假设其中只有t-1属于A0，那么剩下的都在与B1一一对应，有相同数目顶点的A1中。

我们对复现共增加了2/Δ次(3.11)，(Δ-2)/Δ次(3.10)，和1次（3.2）:

ρΔ ≤ 1+2vt·2/Δ+(Δ-2)ρΔ-2/Δ

因为ρ1=ρ2=1，此式可化简为：

ρΔ ≤ (Δ+2)/4+vs/2(Δ为偶数)

ρΔ ≤ (Δ+2+1/3)/4+vs/2(Δ为偶数)

**参考文献**

2. P.Berman and M.Fürer. Approximating maximum independent set in bounded degree graphs. SODA‘94, 365-371.





9. M. M. Halldórsson and J. Radhakrishnan. Improved approximations of independent sets in bounded-degree graphs. SWAT ’94,195-206