最小生成树问题的拓展

安徽省芜湖一中　汪汀

摘要 本文主要论述最小生成树问题中的两类拓展——最小度限制生成树和次小生成树。首先分别介绍了这两类拓展问题的模型，然后提出了求解这两类问题的算法，最后，通过一些例子分析其在实际问题中的应用。

关键字 生成树 拓展 度限制

正文

最小生成树是信息学竞赛中的经典问题，但近年来，竞赛中的题目不再局限于这类经典模型，难度大大增加。为解决这些问题，我们必须对这些经典模型加以拓展。拓展的类型很多，本文主要论述其中的两类——最小度限制生成树和次小生成树。

1 最小生成树

1.1 最小生成树的定义

设 G=(V,E, ω)是连通的无向图，G 中权值和最小的生成树称为最小生成树。

1.2 求解最小生成树的算法

求最小生成树，比较常用的算法有 Prim 算法和 Kruskal 算法。前者借助 Fibonacci 堆可以使复杂度降为 $O(V\log_2 V+E)$，后者一般应用于稀疏图，其时间复杂度为 $O(E\log_2 V)$。

2、最小度限制生成树

2.1、最小度限制生成树的定义

对于一个加权的无向图，存在一些满足下面性质的生成树：某个特殊的结点的度等于一个指定的数值。最小度限制生成树就是满足此性质且权值和最小的一棵生成树。

把它抽象成数学模型就是：

设 G=(V,E, ω)是连通的无向图，$v_0 \in V$ 是特别指定的一个顶点,k 为给定的一个正整数。

如果 T 是 G 的一个生成树且 $d_T(v_0)=k$，则称 T 为 G 的 k 度限制生成树。G 中权值和最小的 k 度限制生成树称为 G 的最小 k 度生成树。

2.2、求解最小度限制生成树的算法

约定：T 为图 G 的一个生成树，T+a-b 记作(+a,-b)，如果 T+a-b 仍然是一个生成树，则称(+a,-b)是 T 的一个可行交换。

引理 1：设 T_1,T_2 是图 G 的两个不同的生成树，$E(T_1)\backslash E(T_2)=\{a_1,a_2,\cdots\cdots,a_n\}$,$E(T_2)\backslash E(T_1)=\{b_1,b_2,\cdots\cdots,b_n\}$,则存在一个排序 $b_{i1},b_{i2},\cdots\cdots,b_{in}$,使得 $T_2+e_j-f_{ij}$ (j=1,2,……,n)仍然是 G 的生成树。

定理 1：设 T 是 G 的 k 度限制生成树，则 T 是 G 的最小 k 度限制生成树当且仅当下面三个条件同时成立：

Ⅰ 对于 G 中任何两条与 v_0 关联的边所产生的 T 的可行交换都是不可改进的。

Ⅱ 对于 G 中任何两条与 v_0 不关联的边所产生的 T 的可行交换都是不可改进的。

Ⅲ 对于 T 的任何两个可行交换$(+a_1,-b_1)$和$(+a_2,-b_2)$，若 a_1,b_2 与 v_0 关联，b_1,a_2 不于 v_0 关联，则有$\omega(b_1)+\omega(b_2)\leqslant\omega(a_1)+\omega(a_2)$

证明：⑴必要性

设 T 是最小 k 度限制生成树，则Ⅰ,Ⅱ显然成立。 以下证明 Ⅲ：由Ⅰ,Ⅱ可知如果$(+a_1,-b_2)$和$(+a_2,-b_1)$都是 T 的可行交换，则有$\omega(b_2)\leqslant\omega(a_1)$, $\omega(b_1)\leqslant\omega(a_2)$，故$\omega(b_1)+\omega(b_2)\leqslant\omega(a_1)+\omega(a_2)$; 否则，或者$(+a_1$,-b2)或者$(+a_2,-b_1)$不是 T 的可行交换，根据引理 1，$T^{'}=T+\{a_1,a_2\}-\{b_1,b_2\}$仍然是 T 的 k 度限制生成树，则$\omega(T)\leqslant\omega(T^{'})$,故$\omega(b_1)+\omega(b_2)\leqslant\omega(a_1)+\omega(a_2)$。

⑵充分性

设 T 是 k 度限制生成树且满足Ⅰ,Ⅱ, Ⅲ,假如有另一个 k 度限制生成树 T’，ω(T’)<ω(T)，设

$E(T')\backslash E(T)=\{a_1,a_2,\cdots\cdots, a_n\}$

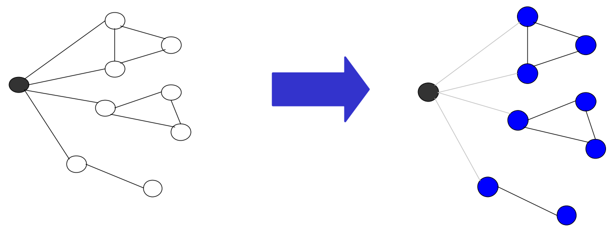
$E(T)\backslash E(T')=\{b_1,b_2,\cdots\cdots, b_n\}$

显然有$\sum\omega(a_i)<\sum\omega(b_i)$，根据引理 1，存在一个排列 $b_{1'},b_{2'},\cdots\cdots,b_{n'}$，满足 $T+a_i-b_{i'}$ 仍然是 G 的生成树。由ω(T’)<ω(T)得$\sum(\omega(b_{i'})-\omega(a_i))>0$，因而，在 T 的这 n 个可行交换中，一定存在某个可以改进的交换（$+a_i,-b_{i'}$）。由于 T 满足Ⅰ,Ⅱ，则 $a_i,b_{i'}$若同时与 v_0 关联或不关联都是不可改进的。也就是说，a_i 和 $b_{i'}$中必定恰好有一个不与 v_0 关联。不妨设 a_i 与 v_0 无关联，因为 $D_{T'}(v_0)$也等于 k，所以必存在另一个交换$(+a_j,-b_{j'})$,满足 a_j 与 v_0 关联，bj’与 v_0 无关联，且$(\omega(b_{i'})-\omega(a_i))+(\omega(b_{j'})-\omega(a_j))>0$,此与Ⅲ矛盾。因此，T’是不存在的，即 T 是 G 的最小 k 度限制生成树。

定理 2：设 T 是 G 的最小 k 度限制生成树，E_0 是 G 中与 v_0 有关联的边的集合，$E_1=E_0\backslash E(T)$，$E_2=E(T)\backslash E_0$，$A=\{(+a,-b)|\ a\in E_1,b\in E_2\}$，设$\omega(a')-\omega(b')=\min\{\omega(a)-\omega(b)|\ (+a,-b)\in A\}$，则 T+a’-b’是 G 的一个最小 k+1 度限制生成树。

如何求最小 k 度限制生成树呢？

首先考虑边界情况。先求出问题有解时 k 的最小值：把 v_0 点从图中删去后，图中可能会出



现 m 个连通分量，而这 m 个连通分量必须通过 v_0 来连接，所以，在图 G 的所有生成树中 $d_T(v_0)\geqslant m$。也就是说，当 $k<m$ 时，问题无解。

再根据上述定理，得出算法的框架：

1 先求出最小 m 度限制生成树； 2 由最小 m 度限制生成树得到最小 m+1 度限制生成树; 3 当 $d_T(v_0)=k$ 时停止。

下面分别考虑每一步

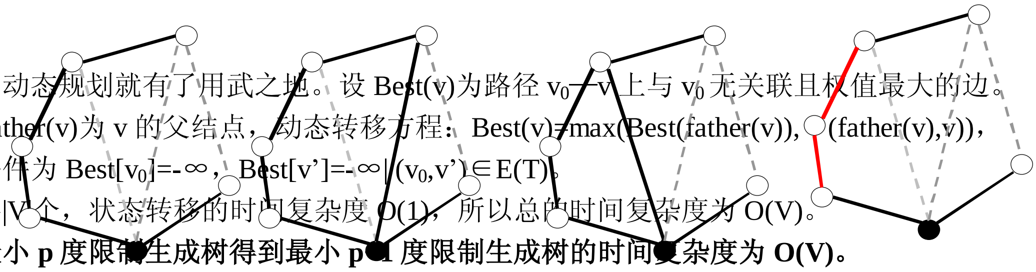
首先，将 v_0 和与之关联的边分别从图中删去，此时的图可能不再连通，对各个连通分量，分别求最小生成树。接着，对于每个连通分量 V’，求一点 v_1，$v_1\in V'$，且 $\omega(v_0,v_1)=\min\{\omega(v_0,v')|v'\in V'\}$，则该连通分量通过边$(v_1,v_0)$与 v_0 相连。于是，我们就得到了一个 m 度限制生成树，不难证明，这就是最小 m 度限制生成树。

这一步的时间复杂度为 $O(V\log_2 V+E)$

我们所求的树是无根树，为了解题的简便，把该树转化成以 v_0 为根的有根树。

假设已经得到了最小 p 度限制生成树，如何求最小 p+1 度限制生成树呢？

根据定理 2，最小 p+1 度限制生成树肯定是由最小 p 度限制生成树经过一次可行交换$(+a_1,-b_1)$得到的。我们自然就有了一个最基本的想法——枚举！但是，简单的枚举，时间复杂度高达 $O(E^2)$，显然是不能接受的。深入思考不难发现，任意可行的交换，必定是一条边跟 v_0 关联，另一条与 v_0 无关，所以，只要先枚举与 v_0 关联的边，再枚举另一条边，然后判断该交换是否可行，最后在所有可行交换中取最优值即可。于是时间复杂度降到了 O(VE)，但这仍然不能令人满意。进一步分析，在原先的树中加入一条与 v_0 相关联的边后，必定形成一个环。若想得到一棵 p+1 度限制生成树，需删去一条在环上的且与 v_0 无关联的边。删去的边的权值越大，则所得到的生成树的权值和就越小。如果每添加一条边，都需要对环上的边一一枚举，时间复杂度将比较高，因为有不少边重复统计多次（下图中红色的边统计了多次）。



这里，动态规划就有了用武之地。设 Best(v)为路径 v_0—v 上与 v_0 无关联且权值最大的边。定义 father(v)为 v 的父结点，动态转移方程：Best(v)=max(Best(father(v)), ω(father(v),v))，边界条件为 Best[v_0]=-∞，Best[v’]=-∞|(v_0,v’)∈E(T)。
状态共|V|个，状态转移的时间复杂度 O(1)，所以总的时间复杂度为 O(V)。
故由最小 p 度限制生成树得到最小 p+1 度限制生成树的时间复杂度为 O(V)。

综上，求最小 k 度限制生成树算法总的时间复杂度为 O(V$\log_2$V+E+kV)。

3、次小生成树

3.1、次小生成树的定义

设 G=(V,E,w)是连通的无向图，T 是图 G 的一个最小生成树。如果有另一棵树 T_1，满足不存在树 T’, ω（T’）<ω（T_1），则称 T_1 是图 G 的次小生成树。

3.2、求解次小生成树的算法

约定：由 T 进行一次可行交换得到的新的生成树所组成的集合，称为树 T 的邻集,记为 N(T)。

定理 3：设 T 是图 G 的最小生成树，如果 T_1 满足ω(T_1)=min{ω(T’)| T’∈N(T)},则 T_1 是 G 的次小生成树。

证明：如果 T_1 不是 G 的次小生成树，那么必定存在另一个生成树 T’,T’=T 使得ω(T)≤ω(T’)<ω(T_1)，由 T_1 的定义式知 T 不属于 N(T)，则
E(T’)\E(T)={a_1,a_2[1],……,a_t}，E(T)\E(T’)={b_1,b_2,……,b_t}，其中 t≥2。根据引理 1 知，存在一个排列 b_{i1},b_{i2},……,b_{it}，使得 T+a_j-b_{ij} 仍然是 G 的生成树，且均属于 N(T),所以ω(a_j)≥ω(b_{ij})，所以ω(T’)≥ω(T+a_j-b_{ij})≥ω(T_1)，故矛盾。所以 T_1 是图 G 的次小生成树。

通过上述定理，我们就有了解决次小生成树问题的基本思路。
首先先求该图的最小生成树 T。**时间复杂度 O(V$\log_2$V+E)**
然后，求 T 的邻集中权值和最小的生成树，即图 G 的次小生成树。
如果只是简单的枚举，复杂度很高。首先枚举两条边的复杂度是 O(VE)，再判断该交换是否可行的复杂度是 O(V)，则总的时间复杂度是 O(V^2E)。这样的算法显得很盲目。经过简单的分析不难发现，每加入一条不在树上的边，总能形成一个环，只有删去环上的一条边，才能保证交换后仍然是生成树，而删去边的权值越大，新得到的生成树的权值和越小。我们可以以此将复杂度降为 O(VE)。这已经前进了一大步，但仍不够好。
回顾上一个模型——最小度限制生成树，我们也曾面临过类似的问题，并且最终采用动态规划的方法避免了重复计算，使得复杂度大大降低。对于本题，我们可以采用类似的思想。首先做一步预处理，求出树上每两个结点之间的路径上的权值最大的边，然后，枚举图中不在

树上的边，有了刚才的预处理，我们就可以用 O(1)的时间得到形成的环上的权值最大的边。

如何预处理呢？因为这是一棵树，所以并不需要什么高深的算法，只要简单的 BFS 即可。预处理所要的时间复杂度为 $O(V^2)$。
这样，这一步时间复杂度降为 $O(V^2)$。

综上所述，次小生成树的时间复杂度为 $O(V^2)$。

4、实际问题中的应用
4.1 野餐计划

矮人虽小却喜欢乘坐巨大的轿车，轿车大到可以装下无论多少矮人。某天，N(N≤20)个矮人打算到野外聚餐。为了集中到聚餐地点，矮人 A 要么开车到矮人 B 家中，留下自己的轿车在矮人 B 家，然后乘坐 B 的轿车同行；要么直接开车到聚餐地点，并将车停放在聚餐地。

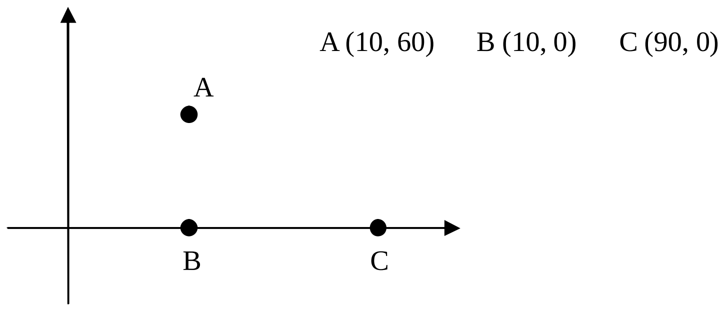
虽然矮人的家很大，可以停放无数量轿车，但是聚餐地点却最多只能停放 K 辆轿车。现在给你一张加权无向图，它描述了 N 个矮人的家和聚餐地点，要你求出所有矮人开车的最短总路程。

[解答]这是一个比较明显的度限制生成树的模型，可以把矮人的家和聚餐地看成图上的点，两个矮人家之间的距离看成一条带权的无向边，聚餐地为有度限制的点。需要注意的是，本题是求度不超过 k 的最小生成树，不过这并没有带来更大的难度，因为从算法的流程来看我们很容易得到度不超过 k 的所有最小度限制生成树。

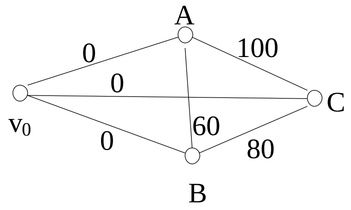
4.2 通讯线路

某地区共有 n 座村庄($1 \le n \le 5000$)，每座村庄的坐标用一对整数(x, y)表示，其中 $0 \le x, y \le 10000$。为了加强联系，决定在村庄之间建立通讯网络。通讯工具可以是需要铺设的普通线路，也可以是卫星设备。卫星设备数量有限，只能给一部分村庄配备。拥有卫星设备的两座村庄无论相距多远都可以直接通讯。而互相间铺设了线路的村庄也可以通讯。现在有 k 台（$0 \le k \le 100$）卫星设备，请你编一个程序，计算出应该如何分配这 k 台卫星设备，才能使铺设线路最短，并保证每两座村庄之间都可以直接或间接地通讯。

[解答]首先构造图，把村庄作为图中的点，村庄间的距离作为边。
如果没有或只有一台卫星设备，就可以直接用最小生成树来解决。卫星设备的作用实际就是连接 k 个点且代价为 0，不妨设一个虚点 v_0，v_0 与原图中的每一个点连接一条代价为 0 的边，v_0 的度限制为 k，再套用度限制生成树的算法即可。
例如下图：



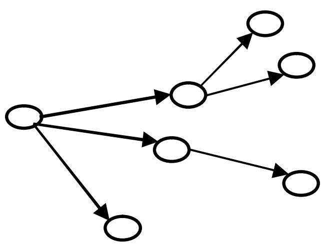
则新构造的图为：



其中，$D_T(v_0)=k$

4.3 秘密的牛奶运输

Farmer John 要把他的牛奶运输到各个销售点。运输过程中，可以先把牛奶运输到一些销售点，再由这些销售点分别运输到其他销售点。



运输的总距离越小，运输的成本也就越低。低成本的运输是 Farmer John 所希望的。不过，他并不想让他的竞争对手知道他具体的运输方案，所以他希望采用费用第二小的运输方案而不是最小的。现在请你帮忙找到该运输方案。

[解答]本题是一个典型的求次小生成树的模型，可以把销售点看成图中的点，每两点间有一条加权的无向边，边的权值为销售点间的距离。那么，直接套用上文所讲述的求次小生成树的算法即可

5、结语

本文主要论述最小生成树问题的两个拓展——度限制生成树和 k 小生成树。其实最小生成树问题的拓展是多种多样的，并非只有本文所提到的两种。当然，不仅仅是最小生成树，

其他经典模型亦是如此。这就需要我们在解决实际问题中，不能拘泥于经典模型，要因“题”制宜，适当地对经典模型加以拓展，建立出符合题目本身特点的模型。

但是，这并不是说，经典模型已经被淘汰。因为一切拓展都是建立在原模型的基础上的，两者之间有着密切的联系。这就需要我们一方面熟练掌握各种经典模型；另一方面，根据实际情况，灵活运用，大胆创新。只有这样，才能在难度日益增加的信息学竞赛中，始终立于不败之地。

参考文献：

[1] 网络算法与复杂性理论 谢政 李建平著

[2] 数据结构（第二版） 严蔚敏 吴伟民著

[3] Introduction to Algorithms, Second Edition Thomas H. Cormen Charles E. Leiserson Ronald L. Rivest Clifford Stein