2022 年第三届 ACM 中国-国际并行计算挑战赛总决赛任务

大规模电源网格系统谱图稀疏化 并行优化

一、 赛题简介

谱图稀疏化算法就是对于给定的一个连通图，求得其某个连通子图，使得原图和子图之间的相对条件数尽可能小。这样的子图便可以作为用 PCG（预处理共轭梯度算法）等迭代算法求解大规模电源网格系统时所需的预条件子，能使之更好更快的收敛。在各种谱图稀疏化算法中，比较常见的有 feGRASS，他是由 Zhuo Feng 等人于 2020 年提出的一种比较简单高效的谱稀疏化算法。它的时间复杂度为$o(m\log m)$，空间复杂度为$o(m)$，m 表示边数。

feGRASS 的基本思想是先对原图做一些简单的预处理后，通过改进的克鲁斯卡尔算法，获得具有一定特征的生成树。再对生成树之外的边的一些属性做计算，根据这些属性，按照特定的标准将一定数量的树外的边添加修复在树上，从而形成目标连通子图。

Algorithm 1 Construction of a MEWST

Input: Graph $G = (V, E, w)$.

Output: MEWST T.

1: Find the vertex with largest volume r . Run breath-first search from r to obtain the unweighted distances between it and all other vertices.

2: Compute the effective weights, via like (19), to get a graph $G' = (V, E, W_{eff})$.

3: Run the Kruskal algorithm on G' to get MEWST T.

其中，W_{eff}的定义为：$W_{\text{eff}}(e) = w_{i,j} \times \dfrac{\log(\max\{\deg(i), \deg(j)\})}{\text{dist}(r, i) + \text{dist}(r, j)}$。$w_{i,j}$是边的权重，$deg(i)$是$i$点的度，$dist(r,i)$是第$i$个点和根结点$r$之间的无权重距离。

Algorithm 2 feGRASS: Fast and Effective Graph Spectral Sparsification

Input: Graph $G = (V, E, w)$, the number of edges added to the spanning tree of G for producing the sparse graph: α.

Output: Sparse graph P, which is spectrally similar to G.

1: Run Algorithm 1 to obtain MEWST T. Set $P = T$.
2: Compute effective resistances R_e of off-tree edges in T.
3: Sort off-tree edges by $w(e)R_e$ in descending order to get an edge list *OffTreeEgdes*.
4: **for** $k = 1$ to $|E| - |V| + 1$ **do**
5: **if** α edges have been added into P **then**
6: Break.
7: **end if**
8: Get edge e =*OffTreeEgdes*[k].
9: **if** $e = (i, j)$ is not marked **then**
10: Add e into P.
11: Compute high-voltage vertices and low-voltage vertices using β-layer BFS from i and j respectively. Mark the off-tree edges similar to e, based on i, j, and the reached vertices in BFS.
12: **end if**
13: **end for**

二、 赛题说明

1. 源码包包括以下文件：
   a. main.cpp 源代码文件；
   b. Eigen 头文件依赖库；
   c. xxx.mtx 输入数据文件；
   d. refer-xxx.txt 参照结果文件。
2. 程序使用方法：
   a. 源码包位置：/public1/soft/IPCC/2022/final/feGRASS.tar
   b. 参考编译器：gcc 8.1.0
   c. 参考编译环境：module load gcc/8.1.0
   d. 基准编译命令：g++ -O3 main.cpp -o main
   e. 集群参考运行命令：srun -p IPCC -N 1 ./main byn1.mtx
3. 以输出的 result.txt 结果文件为评判标准，按照原代码顺序输出生成树的边。结果文件须与参照文件一致，通过 diff result.txt <refer_file> 验证结果准确性。
4. 总决赛阶段发布 5 个数据集：
   a. byn1.mtx 小规模数据，随赛题发布，用于代码调试，**不计入成绩；**
   b. byn.mtx 小规模数据，随赛题发布，可用于热点分析和代码调试，**不计入成绩；**
   c. g-3989-158400.mtx 数据随赛题发布，计入决赛上机成绩，占 1/3；
   d. g-6997-83688.mtx 数据随赛题发布，计入决赛上机成绩，占 1/3；
   e. 第 5 组数据决赛现场早 8:30 发布，计入决赛上机成绩，占 1/3。
5. 在计入成绩的各数据集运行时，必须使用相同运行方式、规模。
6. 计时从读取文件后开始，到写入结果文件前结束，不可修改计时函数位置，**即除数据读写外，其他操作均需计时。**
7. 输入文件 g-3989-158400.mtx，g-6997-83688.mtx 等不可修改。
8. 可以修改数据结构，但**不可使用降低精度或者混合精度优化。**
9. 不可针对数据集本身参数做非通用性优化，包括但不限于诸如：
   a. if param_a == 0 : call func_0();
      else if param_a == 1 : call func_1();
      else : call func();
      ……
   b. if elements > 10000： call MPI_func();
      else : call OpenMP_func();
10. 参赛队员可自行更改编译方式，但需要留存脚本文件或 Makefile 文件。

参考文献：

1）Zhiqiang Liu, Wenjian Yu, Zhuo Feng:feGRASS: Fast and Effective Graph Spectral Sparsification for Scalable Power Grid Analysis. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 41(3): 681-694 (2022).

2）Zhuo Feng，“GRASS: graph spectral sparsifification leveraging scalable spectral perturbation analysis,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 39, no. 12, pp. 4944–4957, 2020.

3）Zhiqiang Liu, Wenjian Yu:pGRASS-Solver: A Parallel Iterative Solver for Scalable Power Grid Analysis Based on Graph Spectral Sparsification. ICCAD 2021: 1-9.

三、 作品内容及要求

1. 优化版源代码

（1） 包含 Makefile，可进行重新编译，并且能够正确生成可执行文件。

（2） 不涉及版权问题，大赛组委会不负责保障源代码安全。

2. 性能优化过程记录表（模板请见附件：1）

3. 技术报告 PPT（模板请见附件：2）及技术报告优化结果页（模板请见附件：3）

（1） 应用程序运行的硬件环境和软件环境，其中软件环境至少包括操作系统、并行环境、相关依赖软件、所运行的应用负载等。

（2） 提供参赛应用程序的代码结构，从设计思路到主要流程设计及主要功能模块。

（3） 详细介绍参赛应用程序中采用的优化方法，基于优化方法达到的优化结果和性能指标。

（4） 详细描述程序运行结果。

四、 竞赛平台

北京超级云计算中心（https://cloud.blsc.cn/）A 分区

五、 竞赛形式与规则

1. 总决赛时间：9 月 22 日-24 日（CCF HPC China 2022 山东-济南 ）
2. 总决赛分为上机及答辩两个环节，所有选手需线下参与相关环节，具体时间安排以《参赛指南》为准。
3. 总决赛环节中，参赛队需于组委会指定平台（https://cloud.blsc.cn/）运行决赛程序。组委会将以程序运行的时间作为上机成绩评分依据。
4. 决赛成绩中，上机成绩多个算例合计占比 80%，技术报告 PPT 答辩占比 20%。注意：，上机成绩以决赛现场提交作品成绩为准，初赛成绩不计入总决赛。
5. 如参赛队发生任何学术不端、违反组委会规定的行为，组委会有权取消参赛队参赛资格，并视情况向所

在单位通报。

六、 联系我们

1. 通知动态：http://www.paraedu.org.cn/
2. 微信：北京超级云计算中心（ID：BJBLSC）
3. 组委会：18310726311 余老师（QQ916034114）
4. 竞赛 QQ 群：1046805935（学生/参赛选手）；1095416620（指导老师）
5. 邮箱：ACM_IPCC@163.com

注意：

1. **代码保存于竞赛服务器，无需下载本地后上传。**
2. 以上内容最终解释权归 ACM 中国-国际并行计算挑战赛组委会所有