CS 430 Project Report

Kai Yao A20347045

Ke Yao A20359070

Problem Definition

Algorithms, implementation and asymptotic behavior for finding spanning trees.

Description of problem

We are going to choose “finding spanning trees” as our project, there are two algorithms for this project. One is Kruskal algorithm, the other is Prim algorithm.

Kruskal algorithm can be described as following:

Suppose there is a graph G, the edge is represented by e and vertex is represented by v, the minimum spanning tree is T. each edge have start vertex and end vertex and its weight.

1. Sort all the edge in increasing order based on its weight.
2. T ←φ, edge ← 0.
3. While(edge < ∣T∣- 1)
4. Pick smallest edge ei.
5. If ei do not form a cycle in T add ei to T.
6. edge++

From this algorithm we can see that we have multiple ways to determine if one edge form a cycle or not. One way is called union set, another way is called DFS.

In union set basically we give each vertex a different set, when picking an edge we see if this edge’s two vertices are in the same set, if it is not, pick this edge, else ignore.

e.start means start vertex, e.end means end vertex, just for clarity.

1. Pick one edge ei
2. Find ei start vertex, and ei end vertex in which set.
3. If two vertices in this edge are not in any set, union them. (do not have a cycle)
4. else If e.start in one set, e.end is not in any set, union them. (do not have a cycle)
5. else If e.end in one set, e.start is not in any set, union them. (do not have a cycle)
6. else If e.start and e.end are in different set, union them. (do not have a cycle)
7. else e.start and e.end are in the same set (have a cycle).

In DFS when picking an edge, we do DFS from any vertex in this edge based on the edges has been choose and connected to this vertex. if we can find another vertex in this edge by DFS, then ignore this edge else pick.

1. Pick an edge ei
2. start ← e.start, destination← e.end
3. flag←DFS(start, destination, T)
4. if(flag)
5. have a cycle
6. else
7. do not have a cycle

DFS:

1. DFS(start, destination, T)
2. mark start
3. if start == destination
4. return true
5. for each adjacent vertex of start w in T
6. if w is unmarked
7. DFS(start, destination, T)
8. return false

Implementation details

Time complexity for Kruskal algorithm with union set

Because we have to do sort based on edges, so the time complexity is at least ElogE.

For each edge∣E∣(at most ∣V∣- 1), we determine if it form a cycle, time complexity is V + ElogV

Because for each edge we have to find two vertices’ set in heap O(2logV), this takes logV and if they are in different set, we have to union them O(1),( number of loop for union set is V)

So the total time complexity is O(ElogE + V + ElogV)

Time complexity for Kruskal algorithm with DSF

Because we have to do sort based on edges, so the time complexity is at least ElogE.

For each edge(at most V- 1), we determine if it form a cycle, time complexity is (V – 1)(V + E)

Because in each vertex, you have to do a DFS, whose time complexity is V + E, in this case should be V + V

So the total time complexity is O(ElogE + V2)

In this project, we use adjacent list to generate vertices and edges.

For star graph, we just generate edges from first vertex to all the rest edges.

For Line graph, we generate edges from 1 to 2, 2 to 3, 3 to 4 and so on so forth. Except for last vertex. It doesn’t need an edge.

For random graph, we generate edges for each vertex, each vertex’s may have 1-10(inclusive) number of edges, which is generated randomly.

Kruskal algorithm Result:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Vertices | Edges | Union set | DFS |
| Star | 1000 | 999 | 40ms | 2597ms |
| Line | 1000 | 999 | 64ms | 540ms |
| Radom | 1000 | 5518 | 15ms | 3422ms |

Prim algorithm Result:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Vertices | Edges | ArrayList | PriorityQueue |
| Star | 1000 | 999 | 2757ms | 1756ms |
| Line | 1000 | 999 | 26ms | 1677ms |
| Radom | 1000 | 5518 | 10ms | 485ms |