COMP251: Minimum Spanning Trees

Jérôme Waldispühl
School of Computer Science
McGill University

Based on (Cormen et al., 2002)
Based on slides from D. Plaisted (UNC)
Minimum Spanning Tree (Example)

- A town has a set of houses and a set of roads.
- A road connects 2 and only 2 houses.
- A road connecting houses $u$ and $v$ has a repair cost $w(u, v)$.

**Goal:** Repair enough (and no more) roads such that:

1. everyone stays connected: can reach every house from all other houses, and
2. total repair cost is minimum.
Model as graph

- Undirected graph $G = (V, E)$.
- **Weight** $w(u, v)$ on each edge $(u, v) \in E$.
- Find $T \subseteq E$ such that:
  1. $T$ connects all vertices ($T$ is a *spanning tree*),
  2. $w(T) = \sum_{(u,v) \in T} w(u, v)$ is minimized.
Minimum Spanning Tree (MST)

- It has $|V| - 1$ edges.
- It has no cycles.
- It might not be unique.
Generic Algorithm

- Initially, A has no edges.
- Add edges to A and maintain the **loop invariant**: “A is a subset of some MST”.

```
A ← ∅;
while A is not a spanning tree do
    find a edge (u, v) that is safe for A;
    A ← A ∪ {(u, v)}
return A
```

- **Initialization**: The empty set trivially satisfies the loop invariant.
- **Maintenance**: We add only safe edges, A remains a subset of some MST.
- **Termination**: All edges added to A are in an MST, so when we stop, A is a spanning tree that is also an MST.
A cut respects $A$ if and only if no edge in $A$ crosses the cut.

A light edge crossing cut (may not be unique)

This edge crosses the cut. (one endpoint is in $S$ and the other is in $V - S$.)

A cut partitions vertices into disjoint sets, $S$ and $V - S$. 
What is a safe edge?

Intuitively: Is \((c,f)\) safe when \(A=\emptyset\)?

- Let \(S\) be any set of vertices including \(c\) but not \(f\).
- There has to be one edge (at least) that connects \(S\) with \(V - S\).
- Why not choosing the one with the minimum weight?
**Theorem 1:** Let $(S, V\setminus S)$ be any cut that respects $A$, and let $(u, v)$ be a light edge crossing $(S, V\setminus S)$. Then, $(u, v)$ is safe for $A$.

**Proof:**
Let $T$ be a MST that includes $A$.

**Case 1:** $(u, v)$ in $T$. We’re done.

**Case 2:** $(u, v)$ not in $T$. We have the following:

Let $T' = T - \{(x, y)\} \cup \{(u, v)\}$. Because $(u, v)$ is light for cut, $w(u, v) \leq w(x, y)$. Thus, $w(T') = w(T) - w(x, y) + w(u, v) \leq w(T)$. Hence, $T'$ is also a MST. So, $(u, v)$ is safe for $A$. 

We show edges in $T$.
Corollary

In general, A will consist of several connected components.

**Corollary:** If \((u, v)\) is a light edge connecting one CC in \((V, A)\) to another CC in \((V, A)\), then \((u, v)\) is safe for A.
Kruskal’s Algorithm

1. Starts with each vertex in its own component.
2. Repeatedly merges two components into one by choosing a light edge that connects them (i.e., a light edge crossing the cut between them).
3. Scans the set of edges in monotonically increasing order by weight.
4. Uses a **disjoint-set data structure** to determine whether an edge connects vertices in different components.
Example
Example

Graph with labeled nodes and edges:
- Nodes: a, b, c, d, e, f, g, h, i
- Edges: 
  - a to b: 10
  - b to c: 9
  - c to d: 8
  - d to e: 7
  - e to f: 3
  - f to g: 5
  - g to i: 2
  - c to f: 1
  - f to h: 6

Weights for the edges:
- a to b: 10
- b to c: 9
- c to d: 8
- d to e: 7
- e to f: 3
- f to g: 5
- g to i: 2
- c to f: 1
- f to h: 6
Example

Reject!
Example
Example
Example
Example

Graph with labeled vertices and edges:

- Vertices: a, b, c, d, e, f, g, h, i
- Edges with weights:
  - ab: 10
  - bc: 9
  - cd: 8
  - de: 7
  - ef: 3
  - fg: 5
  - gh: 6
  - hi: 11
  - ad: 8
  - bh: 2
  - ce: 3
  - de: 7
  - ef: 3
  - gh: 6
  - ih: 11

Weights are shown as numbers adjacent to the edges.
Kruskal’s complexity

- Initialize $A$: $O(1)$
- First **for** loop: $|V|$ MAKE-SETs
- Sort $E$: $O(E \lg E)$
- Second **for** loop: $O(E)$ FIND-SETs and UNIONs

Assuming union by rank and path compression:
$O((V + E)\alpha(V)) + O(E \lg E)$

- Since $G$ is connected, $|E| \geq |V| - 1 \Rightarrow O(E \alpha(V)) + O(E \lg E)$.
- $\alpha(|V|) = O(\lg V) = O(\lg E)$.
- Therefore, total time is $O(E \lg E)$.
- $|E| \leq |V|^2 \Rightarrow \lg |E| = O(2\lg V) = O(\lg V)$.

$\Rightarrow O(E \lg V)$ time
Prim’s Algorithm

1. Builds **one tree**, so $A$ is always a tree.
2. Starts from an arbitrary “root” $r$.
3. At each step, **adds a light edge** crossing cut $(V_A, V - V_A)$ to $A$.
   - Where $V_A$ = vertices that $A$ is incident on.
Intuition behind Prim’s Algorithm

• Consider the set of vertices $S$ currently part of the tree, and its complement ($V-S$). We have a cut of the graph and the current set of tree edges $A$ is respected by this cut.

• Which edge should we add next? *Light edge!*
Finding a light edge

1. Uses a priority queue $Q$ to find a light edge quickly.
2. Each object in $Q$ is a vertex in $V - V_A$.
3. Key of $v$ has minimum weight of any edge $(u, v)$, where $u \in V_A$.
4. Then the vertex returned by Extract-Min is $v$ such that there exists $u \in V_A$ and $(u, v)$ is light edge crossing $(V_A, V - V_A)$.
5. Key of $v$ is $\infty$ if $v$ is not adjacent to any vertex in $V_A$. 
Basics of Prim’s Algorithm

• It works by adding leaves on at a time to the current tree.
  – Start with the root vertex $r$ (it can be any vertex). At any time, the subset of edges $A$ forms a single tree. $S = \text{vertices of } A$.
  – At each step, a light edge connecting a vertex in $S$ to a vertex in $V - S$ is added to the tree.
  – The tree grows until it spans all the vertices in $V$.

• Implementation Issues:
  – How to update the cut efficiently?
  – How to determine the light edge quickly?
Implementation: Priority Queue

• Priority queue implemented using heap can support the following operations in $O(lg n)$ time:
  – Insert $(Q, u, key)$: Insert $u$ with the key value $key$ in $Q$
  – $u = \text{Extract}_\text{Min}(Q)$: Extract the item with minimum key value in $Q$
  – $\text{Decrease}_\text{Key}(Q, u, \text{new}_\text{key})$: Decrease the value of $u$'s key value to $\text{new}_\text{key}$

• All the vertices that are not in the $S$ (the vertices of the edges in $A$) reside in a priority queue $Q$ based on a $key$ field. When the algorithm terminates, $Q$ is empty. $A = \{(v, \pi[v]): \ v \in V - \{r\}\}$
Prim’s Algorithm

Q := V[G];
for each u ∈ Q do
  key[u] := ∞
  π[u] := Nil;
  Insert(Q,u)
Decrease-Key(Q,r,0);
while Q ≠ ∅ do
  u := Extract-Min(Q);
  for each v ∈ Adj[u] do
    if v ∈ Q ∧ w(u, v) < key[v] :
      π[v] := u;
      Decrease-Key(Q,v,w(u,v));

Notes: (i) A = {(v, π[v]) : v ∈ v - {r} - Q}. (ii) r is the root.
Example of Prim’s Algorithm

Q = a b c d e f
0 ∞ ∞ ∞ ∞ ∞ ∞
Example of Prim’s Algorithm

Q = b d c e f
5 11 ∞ ∞ ∞
Example of Prim’s Algorithm

Q = e c d f
3 7 11 ∞
Example of Prim’s Algorithm

Q = d c f
0 1 2
Example of Prim’s Algorithm

\[ Q = c \ f \]

\[
\begin{array}{c}
\text{a/0} & \text{b/5} & \text{c/1} \\
11 & 3 & 1 \\
5 & 0 & 2 \\
\end{array}
\]

\[
\begin{array}{c}
\text{d/0} & \text{e/3} & \text{f/2} \\
1 & -3 & \\
\end{array}
\]
Example of Prim’s Algorithm

Q = f
-3
Example of Prim’s Algorithm

Q = ∅
Example of Prim’s Algorithm
Correctness of Prim

• Again, show that every edge added is a safe edge for $A$
• Assume $(u, v)$ is next edge to be added to $A$.
• Consider the cut $(A, V-A)$.
  – This cut respects $A$
  – and $(u, v)$ is the light edge across the cut
• Thus, by the Theorem 1, $(u, v)$ is safe.