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)
Problem

• 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”.

```plaintext
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.

**Definitions**

- **Cut** partitions vertices into disjoint sets, $S$ and $V - S$.

- A light edge crossing cut (may not be unique)

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

Theorem 1: Let $(S, V-S)$ be any cut that respects $A$, and let $(u, v)$ be a light edge crossing $(S, V-S)$. Then, $(u, v)$ is safe for $A$.

Let $(x, y)$ crosses cut. 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 $A$. 

---

**Safe edge**

**Theorem 1:** Let $(S, V-S)$ be any cut that respects $A$, and let $(u, v)$ be a light edge crossing $(S, V-S)$. Then, $(u, v)$ is safe for $A$.
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

![Graph Image]
Example
Example

Reject!
Example
Example
Example
Example
Example

Graph with labeled nodes and edges:

- Nodes: a, b, c, d, e, f, g, h, i
- Edges with weights:
  - a to b: 9
  - b to c: 8
  - c to d: 3
  - d to e: 7
  - e to f: 3
  - f to g: 6
  - g to i: 9
  - a to i: 12
  - b to d: 8
  - d to g: 5
  - g to h: 2
  - h to i: 11
Example

Graph with vertices labeled a, b, c, d, e, f, g, h, i and edges labeled with numbers 1 through 12.
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 $V_A = \text{vertices that } A \text{ is incident on.}$
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 $v$ is 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 $V_A$. 
Prim’s Algorithm

\[ Q := V[G]; \]
\[ \text{for each } u \in Q \text{ do} \]
\[ \quad \text{key}[u] := \infty \]
\[ \text{od;} \]
\[ \text{key}[r] := 0; \]
\[ \pi[r] := \text{NIL}; \]
\[ \text{while } Q \neq \emptyset \text{ do} \]
\[ \quad u := \text{Extract-Min}(Q); \]
\[ \quad \text{for each } v \in \text{Adj}[u] \text{ do} \]
\[ \quad \quad \text{if } v \in Q \land w(u, v) < \text{key}[v] : \]
\[ \quad \quad \quad \pi[v] := u; \]
\[ \quad \quad \quad \text{key}[v] := w(u, v) \]

Complexity:
Using binary heaps: \( O(E \lg V) \).
   - Initialization: \( O(V) \).
   - Building initial queue: \( O(V) \).
   - \( V \) Extract-Min: \( O(V \lg V) \).
   - \( E \) Decrease-Key: \( O(E \lg V) \).

Using Fibonacci heaps:
\( O(E + V \lg V) \).

\[ \Leftrightarrow \text{ decrease-key operation} \]

Note: \( A = \{(v, \pi[v]) : v \in V - \{r\} - Q\} \).
Example of Prim’s Algorithm

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

\[ Q = e \ c \ d \ f \]
\[3 \ 7 \ 11 \ \infty\]
Example of Prim’s Algorithm
Example of Prim’s Algorithm

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

\[ Q = f \]
\[ Q = -3 \]
Example of Prim’s Algorithm

Q = ∅
Example of Prim’s Algorithm