Lecture 6: Approximation Algorithms for the TSP  
(Travelling Salesman Problem)\(^1\)

1 Introduction

The Travelling Salesman Problem
Given \( n \) cities, for each pair of distinct cities there is a cost (weight or distance) between them. Compute a simple closed path (i.e. a path that starts and ends at the same city and visits each city exactly once) that minimizes the sum of the edges along the path.

Input: A set of \( n \geq 3 \) cities. To each pair of cities \( (i, j) \) we associate a cost (weight or distance), \( d_{i,j} \).

Output: A closed path or tour of minimum cost, where the cost of the tour is defined as the sum of the weights of the edges in the tour.

1.1 The Metric TSP

In this lecture, we consider a special case of TSP called Metric TSP. Metric TSP is a subcase of TSP where the Triangle Inequality holds. (This is true when the weights on the edges are actual distances).

Definition 1.1.1. The Triangle Inequality:

\[
\forall (i, j, k) \quad d(i, j) \leq d(i, k) + d(k, j)
\]

2 Approximation Algorithms for Metric TSP

Both TSP with arbitrary edge weights, and the special case of Metric TSP are NP-hard problems, that is, there cannot be a polynomial-time algorithm for solving these problems unless P=NP. But, we can use approximation algorithms to get within a certain factor of the optimal answer for the Metric TSP problem.

Let \( T^* \) denote the cost of the minimum weight tour and \( |T^*| \) be the cost of the tour.

Goal 1: A polynomial-time algorithm that outputs a tour of cost \( C \), where \( C \leq 2|T^*| \). Investigated in the previous lecture’s notes.

\(^1\)These notes were revised from scribe notes from Stephanie Tauber in 2002
Goal 2: Christofides’ Algorithm: A polynomial-time algorithm that outputs a tour of cost $C$, where $C \leq \frac{3}{2}|T^*|$. 

3 Graph Properties

Definition 3.0.2. Given a graph, $G$, an Eulerian tour of $G$ is a closed walk that traverses each edge of $G$ exactly once. A Hamiltonian tour of $G$ is a closed walk that visits each node of $G$ exactly once (except for the first and last node). A graph is Eulerian if it permits an Eulerian tour and Hamiltonian if it permits a Hamiltonian tour.

Lemma 3.0.3. Let $G$ be a graph with trail $C$ (a tour that can be formed without repeating an edge), then

1. Construct $G'$ from $G$ by removing all loops from $G$. Then $G$ has no vertices of odd degree iff $G'$ has no vertices of odd degree.

2. Construct $C'$ from $C$ by short-cutting each of the loops in $C$. Then $C$ is an Eulerian tour of $G$ iff $C'$ is an Eulerian tour of $G'$.

Proof. 1. Assume $G$ has no vertices of odd degree and vertex $v$ has a loop, then removing this loop from $G$ results in a new graph $H$ where the degree of vertex $w$, $d(w)$, in graph $H$ equals $d(w)$ in $G$ if $w \neq v$ and $d(w)$ in graph $H$ equals $d(w) - 2$ in $G$ if $w = v$. Thus if $d(w)$ is even for all vertices of $G$, then it is also even for all vertices of $H$. $H$ is still connected as $G$ was connected and removing a loop cannot disconnect $G$ as both endpoints of a loop are the same vertex. The proof in the other direction is similar.

2. Assume $G$ has an Eulerian tour $C$ and vertex $v$ has a loop, then removing this loop from $G$ results in a new graph $H$. Let $C = c_1l_vc_2$, where $c_1$ is the part of $C$ before the loop at $v$ is reached, $l_v$ is the loop at $v$, and $c_2$ is the part of $C$ after the loop at $v$ is reached. Then $D = c_1c_2$ is an Eulerian tour of $H$ as $c_1$ ends at vertex $v$ and $c_2$ begins at vertex $v$, as both endpoints of $l_v$ are $v$, thus $D$ is a tour. Moreover, each edge of $H$ occurs once in $D$ because otherwise, it would not occur in $C$ and then $C$ would not be an Eulerian tour of $G$. The proof in the other direction is similar.

\[\Box\]

Theorem 3.0.4. A connected graph, $G$, is Eulerian iff it has no vertices of odd degree.

Proof. (⇒) Let $G$ be Eulerian and let $C$ be an Eulerian tour of $G$ that starts and ends at vertex $u$, wlog we can assume that $G$ has no loops. Each time a vertex $v \neq u$ occurs at an internal vertex of $C$, two edges incident to $v$ are accounted for, thus all edges incident to
can be paired in this manner as an Eulerian tour contains each edge of $G$ exactly once, thus $d(v)$ is even. Similarly, for $u$, every time $u$ occurs as an internal vertex of $C$, two edges incident to $u$ are accounted for, and the first and last edges of $C$ can be paired, thus all edges incident to $u$ can be paired in this manner as an Eulerian tour contains each edge of $G$ exactly once, thus $d(u)$ is even.

(⇐) Suppose $G$ is a non-Eulerian connected graph with at least 1 edge and no vertices of odd degree, wlog assume that $G$ does not have any loops. Choose such a graph $G$ with as few edges as possible. Since each vertex of $G$ has degree 2, then we can form a closed trail as follows: Let $u$ be a vertex of $G$ and $v_1$ a vertex connected to $u$ by an edge. Start the trail with this edge, as $d(v_1)$ is even there is an unused edge incident to $v_1$. Let the vertex at the other end of this edge be $v_2$, if $v_2 = u$, then we have a trail of positive length, otherwise, continue similarly from $v_2$. Eventually, some $v_i$ will equal $u$ as there are a finite number of edges in the graph and if $v_i \neq u$ then there exists a $v_{i+1}$ reacheable on an unused edge. The previous shows that we can always construct a closed trail with positive length. Let a closed trail of maximum possible length of $G$ be $C$.

Delete the edges of $C$. What remains is a set of connected components, $(C - G)$, where the number of edges of each component is less than the number of edges of $G$. Let $D$ be one of these connected components. As $D$ has less edges than $G$, $D$ is Eulerian by hypothesis, let the Eulerian tour be $C'$. As $G$ was connected, this Eulerian tour can be added to $C$ to create a longer trail and $C \cap C' \neq \emptyset$. Let $v$, a vertex, be in $C \cap C'$, assume wlog that $C$ and $C'$ both start and end at $v$, but then $C$ followed by $C'$ is a longer trail of $G$, so $C$ was not the largest trail, thus by contradiction $G$ has an Eulerian tour.

The second half of this proof hints at a divide and conquer method to find an Eulerian tour on a graph.

Claim 3.0.5. In any graph, the number of vertices of odd degree must be even.

Proof. As the sum of the degrees of all the vertices is twice the number of edges (each end of an edge increases the total degree of the graph by 1),

$$\sum_{v \in V} \deg(v) = 2|E|$$

Moreover, the vertices of the graph can be partitioned into two sets, $V_{\text{odd}}$, the vertices of odd degree and $V_{\text{even}}$, the vertices of even degree.

$$\sum_{v \in V} \deg(v) = \sum_{v \in V_{\text{odd}}} \deg(v) + \sum_{v \in V_{\text{even}}} \deg(v)$$

As the sum of degrees of all vertices is even as it equals twice the number of edges and the sum of degrees of all vertices of even degree is even as the sum of evens is even, then the sum of the degrees of the vertices of odd degree is even. Then as the sum of $k$ odds is even iff $k$ is even, the number of vertices of odd degree is even.

Fact 3.0.6. A minimum weight perfect matching in a weighted complete graph of size $2n$ can be found in polynomial time. [2]
4 Goal 2: $C \leq \frac{3}{2}|T^*|$ (Christofides’ Algorithm)

1. Create a minimum spanning tree on the set of cities, $MST$. From the previous lecture we know that $|MST| \leq |T^*|$. 

2. In this $MST$, by claim 3.0.5 the number of vertices of odd degree is even. Find a minimum-weight perfect matching $M^*$ in the original graph between the vertices that have odd degree in the $MST$ (The perfect matching exists as this is a complete graph).

Claim 4.0.7. $|M^*| \leq \frac{1}{2}|T^*|$

Proof. Let the set $V = \{v_1, \ldots, v_{2k}\}$ be the vertices of $MST$ that have odd degree. Then let $M$ be the Hamiltonian path on $V$ that visits the vertices in the same order that $T^*$ visits the vertices of $V$. Then by the triangle inequality, $|M| \leq |T^*|$. Now number the edges of $M$ in the order that they are reached in the Hamiltonian path. Define two perfect matchings on $V$, $M_1$ and $M_2$, where the odd numbered edges from the labelling are in $M_1$ and the even numbered edges are in $M_2$. (An alternative way of thinking of this is putting every other edge in $M_1$ and the remaining edges in $M_2$) Then $\min\{|M_1|, |M_2|\} \leq \frac{1}{2}|M| \leq \frac{1}{2}|T^*|$. Wlog let $|M_1|$ be the smaller of the two. As $M^*$ is the minimum weight perfect matching on the vertices of odd degree, so $|M^*| \leq |M_1|$, thus

$$|M^*| \leq |M_1| = \min\{|M_1|, |M_2|\} \leq \frac{1}{2}|M| \leq \frac{1}{2}|T^*|$$

3. Consider the subgraph $MST + M^*$, every vertex of this graph has even degree (as $M^*$ adds 1 to the degree of each of the vertices of odd degree of $MST$), thus there exists an Eulerian tour, $E$, of this subgraph with cost $|MST| + |M^*| \leq \frac{3}{2}|T^*|$ since each edge is used exactly once.

4. Write down each vertex the first time it appears in the Eulerian tour $E$ (and thus short-cutting the path by the triangle inequality), creating a salesman tour $E^*$ with cost $\leq \frac{3}{2}|T^*|$.

References
