Paper for USRA 2025 - Studing the Wiener Index for spatially embedded graphs

Atishaya Maharjan

maharjaa@myumanitoba.ca

University of Manitoba— June 9, 2025

Abstract

Include your abstract when you are done here.

1 Introduction

The Wiener index is a topological index that is defined as the sum of distances between all pairs of vertices in a graph. It is widely used in chemistry to predict the properties of molecules based on their structure. In this paper, we will study the Wiener index for spatially embedded graphs, which are graphs that can be drawn in a Euclidean space.

Mathematically, the *Wiener index* of a weighted graph $G = (V, E)$ is the sum of the distances between all pairs of vertices in G:

$$W(G) = \sum_{u,v \in V} \delta_G(u, v)$$

Where $\delta_G(u, v)$ is the weight of the shortest (minimum-weight) path between vertices u and v in the graph G. Formally, the Wiener index was introduced by the chemist Harry Wiener in 1947 [2]. It hand its variations have several applications in chemistry, biology, and network analysis.

> Add citation about the various applications here

2 Preliminaries

Let P be a set of n points in the Euclidean plane $\mathbb{R}^2$. The *Euclidean distance* between two points $p_i, p_j \in P$ is defined as:

$$|p_i p_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Let $G = (P, E)$ be the complete graph on the set of points P, where the edge set E consists of all pairs of points in P. The weight of an edge $(p_i, p_j) \in E$ is defined as the Euclidean distance between the points p_i and p_j and is denoted as $w(p_i, p_j) = |p_i p_j|$.

Let T be a spanning tree of the graph G (That is, it spans all the points in P and is connected). For each pair of points $(p, q) \in P$, let $\delta_T(p, q)$ be the weight of the (unique) shortest path between p and q in the tree T.

Then, we define the *Wiener index* of the tree T as:

$$W(T) = \sum_{p,q \in P} \delta_T(p, q)$$

Finally, we define the total weight of the paths in T from p to every other point $q \in P$ as:

$$\delta_p(T) = \sum_{q \in P} \delta_T(p, q)$$

> Add more details about the Wiener index, its properties, a brief overview about the work done to find the optimal way to compute it for different graphs, and the challenges in computing it for spatially embedded graphs.

> Add a section about our contribution here.

3 Computing the Wiener index for spatially embedded graphs

Abuaffash et al. [1] showed the following theorems:

Theorem 3.1 (Abuaffash et al. 2023). *The spanning tree of P that minimizes the Wiener index is planar.*

Theorem 3.2 (Abuaffash et al. 2023). *There exists a $O(n^4)$ time algorithm that constructs a spanning tree of P that minimizes the Wiener index.*

Theorem 3.3 (Abuaffash et al. 2023). *Given a cost W and a budget B, computing a spanning tree of P whose Wiener index is at most W and whose total weight of the paths from each point to every other point is at most B is weakly NP-hard.*

Theorem 3.4 (Abuaffash et al. 2023). *The Hamiltonian path of P that minimizes the Wiener index is not necessarily planar and computing it is NP-hard.*

We tackle the problem of computing the Wiener index for points in general position as well as give an approximation algorithm for the problem of computing the hamiltonian path of P that minimizes the Wiener index.

3.1 Constructing a spanning tree of P in general position that minimizes the Wiener index

3.2 An approximation algorithm for the Hamiltonian path of P that minimizes the Wiener index

We use Euclidean Minimum Spanning trees to show a lower bound on the Wiener index of a Hamiltonian path of P that minimizes the Wiener index.

The big idea:

1. The EMST seems like a good heuristic/approximation for the Wiener index.

2. Generate the EMST using Kruskal's or Prim's algorithm.

3. Double the edges of the EMST to create an Eulerian circuit.

4. Use the Eulerian circuit to create a Hamiltonian path by performing a depth-first traversal of the circuit.

That's a classic TSP 2-approximation via MST. We could probably use the same structure but analyze it for Wiener index instead of total path length.

Question:

1. Can we show some kind of bound of total bound length VS Wiener index?

2. Can we show that the total path length of the MST is at most some factor larger than the Wiener index of the optimal Hamiltonian path?

Steps to be taken:

- Define P_{MST}: The Hamiltonian path derived from the Euclidean Minimum Spanning Tree of P.

- Let H^* be the optimal Wiener-Minimal Hamiltonian path.

- Show that the Wiener index of P_{MST} is at most a α factor larger than the Wiener index of H^*. Ideally, α is logarithmic or constant.

- Use the triangle inequality and Euclidean metric properties to relate path distance and the Euclidean distance.

- Bound how much "detour" the path introduces in terms of total pairwise distance.

- Verify experimentally that the approximation holds for various point distributions. (TBD ?)

Algorithm

Given a set of points $P = \{p_1, p_2, \ldots, p_n\}$ in the Euclidean plane:

1. Compute the Euclidean Minimum Spanning Tree (EMST) T of the points P using Kruskal's or Prim's algorithm.

2. Perform a Depth-First Search (DFS) traversal on T, starting from an arbitrary root vertex $r \in P$, recording the order in which vertices are first visited.

3. Construct a Hamiltonian path H by listing the vertices in the order they are first encountered during the DFS traversal, skipping any repeated vertices (i.e., shortcutting repeated visits).

Constructing a Hamiltonian path from a Wiener-Minimal Spanning Tree

Given $P \subseteq \mathbb{R}^2$ be a set of n points. Let T be a Wiener-Minimal Spanning Tree of P. Let H be the Hamiltonian path constructed from T by performing a preorder DFS traversal on T.

Let $W(T)$ be the Wiener index of the tree T and $W(H)$ be the Wiener index of the Hamiltonian path H.

Theorem 3.5. *Let T be a spanning tree on a set of n points P, and let H be a Hamiltonian path obtained by performing a depth-first traversal of T and shortcutting repeated visits. Then:*

$$W(H) \leq (n-1) \cdot W(T)$$

Proof. Let $d_T(u, v)$ denote the tree distance between vertices u and v, i.e., the length of the unique path connecting u and v in T. Let $d_H(u, v)$ denote the path distance between u and v along the Hamiltonian path H.

We begin by noting that the Wiener index of any graph G is defined as:

$$W(G) = \sum_{\{u,v\} \subseteq P} d_G(u, v)$$

Hence, we want to bound:

$$W(H) = \sum_{\{u,v\}} d_H(u, v)$$

in terms of $W(T) = \sum_{\{u,v\}} d_T(u, v)$.

The Hamiltonian path H is constructed by performing a DFS traversal of T and shortcutting repeated visits (i.e., only recording the first occurrence of each vertex). The DFS traversal of a tree visits each edge twice (once when descending, once when backtracking), and thus its total traversal length is exactly:

$$L_{\text{DFS}} = 2 \cdot \text{weight}(T)$$

Now consider any pair of vertices $u, v \in P$. In the worst case, u and v lie at opposite ends of the Hamiltonian path H, and the distance between them in H is the sum of up to $n-1$ edges (That is, the diameter of the tree). Since the Hamiltonian path visits each vertex exactly once, we have:

$$d_H(u, v) \leq diam(T) \cdot \max_{e \in T} w(e)$$

Meanwhile, in the tree T, we have:

$$d_T(u, v) \geq \min_{e \in T} w(e)$$

Since H is a traversal of T, any subpath between u and v in H can include at most $n-1$ edges, while $d_T(u, v)$ includes the minimal number of edges between u and v in T.

In the worst case (such as when T is a star graph), the Hamiltonian path must "walk around" the entire outer ring to connect two leaf nodes. This inflates some distances up to $(n-1)$ times their tree values:

$$d_H(u, v) \leq (n-1) \cdot d_T(u, v)$$

Hence, for each pair $\{u, v\}$, the inflation factor is:

$$\alpha_{uv} = \frac{d_H(u, v)}{d_T(u, v)} \leq n - 1$$

Applying this uniformly over all $\binom{n}{2}$ unordered pairs gives:

$$W(H) = \sum_{\{u,v\}} d_H(u, v) \leq (n - 1) \cdot \sum_{\{u,v\}} d_T(u, v) = (n - 1) \cdot W(T)$$

Therefore, we obtain the approximation guarantee:

$$W(H) \leq (n - 1) \cdot W(T)$$

which implies a factor-$O(n)$ approximation for the Wiener index via DFS-induced Hamiltonian paths. □

4 Higher dimensional generalization

5 Hypergraphs and the Wiener index

6 Proving something about Hypertrees (Or other structures) that minimizes the Wiener index

7 Conclusion

In this paper, we studied whatever that we needed to study :).

References

[1] A. Karim Abu-Affash, Paz Carmi, Ori Luwisch, and Joseph S. B. Mitchell. Geometric spanning trees minimizing the wiener index, 2023.

[2] Harry Wiener. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1):17–20, 1947.