arXiv:2303.01096v1 [cs.CG] 2 Mar 2023

使维纳指数最小的几何生成树

A. Karim Abu-Affash[1], Paz Carmi[2], Ori Luwisch[2], 和 Joseph S. B. Mitchell[3]

[1] 以色列沙蒙工程学院软件工程系 abuaa1@sce.ac.il[2] 以色列本-古里安大学计算机科学系 carmip@cs.bgu.ac.il orilu@post.bgu.ac.il[3] 美国石溪大学应用数学与统计学系 joseph.mitchell@stonybrook.edu

摘要。 网络的维纳指数由化学家 Harry Wiener [30], 提出，是网络中所有节点对之间距离的总和。该指数最初用于表示分子中非氢原子的化学图，现已被认为是一种基础且实用的网络描述符。我们研究了在欧几里得空间的点集上构造使维纳指数最小的几何网络这一问题：给定欧几里得空间中的一组 P n 个 $\mathbb{R}^d$点，目标是构造一个覆盖这些点、 P 并满足某些约束的网络，使其在允许的生成网络类别中维纳指数最小。

在这项工作中，我们主要关注作为树的生成网络，并研究平面（$d = 2$）中的问题。我们证明，任何使维纳指数最小化的生成树，其边在平面中都互不相交。随后，我们利用这一事实设计了一种耗时 $O(n^4)$的算法，用于为凸位置点构造维纳指数最小的生成树。我们还证明，在 P上计算一棵维纳指数至多为 W、且总（欧几里得）权重至多为 B的生成树这一问题是 NP 困难的。

在通信网络领域，使维纳指数最小化的树的计算已有研究，这一问题称为最优通信生成树问题。

关键词： 维纳指数 · 最优通信生成树 · 最小路由代价生成树。

1 引言

带权图 $G = (V, E)$的维纳指数是图中每对顶点之间最短路径长度之和，即$\sum_{u, v \in V} \delta_G(u, v)$；其中$\delta_G(u, v)$是 u 与 v 在G中的最短（权重最小）路径的权重。维纳指数由化学家 Harry Wiener 于 1947 [30]提出。维纳指数及其若干变体已应用于化学领域，

例如，用于预测药物的抗菌活性和模拟晶体现象。它还被用于深入了解分子的各种化学和物理性质 [28] ，并将分子的结构与其生物活性相关联 [20]。维纳指数已成为一般科学文化的一部分，至今仍是深入研究的课题 [2,10,12,32]。在化学领域的应用中，维纳指数最常在无权图的背景下进行研究。点间距离之和最小化的研究也自然地出现在网络设计领域；在该领域中，计算维纳指数最小的生成树这一问题称为最优通信生成树（OCST）问题 [15,18]。

给定一个无向图 $G = (V, E)$，以及定义在图 G 的边上的（非负）权函数，用于表示每条边的延迟。图 G 的生成树 T 的路由代价 $c(T)$，等于 T 中每对顶点之间路径的权重（延迟）之和： $c(T) = \sum_{u, v \in V} \delta_T(u, v)$，其中 $\delta_T(u, v)$ 是图 T 中连接 u 和 v 的唯一路径的权重。OCST问题旨在为给定的带权无向图 G 找到一棵路由代价最小的生成树，从而使树中两个随机选取的顶点之间路径的期望代价最小化。Hu [18] 最初提出了 OCST 问题；已知该问题在图上是 NP完全的，即使所有边权均为 1 [19] 也是如此。Wu等人 [31] 针对OCST问题提出了一种多项式时间近似方案（PTAS）。具体而言，他们证明，最优的 k-星形树（内部顶点不超过 k 个的树）可为该问题提供 $\binom{k+3}{k+1}$ 近似，从而得到一种运行时间为 $O\left(n^{2\lceil \frac{2}{\varepsilon} \rceil - 2}\right)$ 的 $(1+\varepsilon)$ 近似算法。

尽管关于维纳指数的研究十分丰富，例如计算特定图或图类的维纳指数及其界限 [16,17,24]，以及推导特殊图类的维纳指数显式公式 [3,23,26,29,30],，但据我们所知，维纳指数在几何环境中尚未受到太多关注。本文研究若干特定几何环境中的维纳指数和最优通信生成树问题，希望让计算几何领域的研究者关注这一重要且应用广泛的指数。

本文贡献与概述。设 P 是平面上的一组 n 点。我们研究如下问题：当底层图是以 P为顶点集的完全图，且边权为相应点之间的欧几里得长度时，求一棵使维纳指数最小的生成树。在第 2 节中，我们证明最优树（即使维纳指数最小的树）没有相交边。作为主要的算法结果，我们在第 3 节给出一个多项式时间算法，在点集 P 处于凸位置时求解该问题；这一结果充分利用了最优树的边不会相交这一结构性结果，从而使我们能够设计动态规划算法来求最优解。随后，在第 4 节中，我们证明“欧几里得维纳指数树问题”是（弱）NP 困难的；该问题要求求出一棵以 P 为顶点集、维纳指数至多为 W且总（欧几里得）权重至多为 B的生成树。最后，在第 5 节中，我们讨论如何寻找维纳指数最小的路径生成 P。

相关工作。与我们研究的问题相关的是最小延迟问题，也称为旅行维修工问题 TRP：求一条路径，从点 s出发，访问所有点，同时使沿路径从 s 到每个其他点的距离之和（即“延迟”）最小（而非维纳指数中所有点对之间的距离）。在欧几里得平面和带权平面图中，TRP（以及由 k 名维修工执行的 k-TRP）存在多项式时间近似方案 [27]。

维纳指数优化也出现在将一个度量空间非收缩地嵌入另一个度量空间（例如直线度量或树度量）的研究背景中，其目的是最小化嵌入的平均失真（定义为新空间中所有点对距离之和除以原空间中所有点对距离之和）。将树度量嵌入直线度量以最小化平均失真是 NP 困难的；有一种基于 k-TRP 的常数因子近似算法，可用于最小化将度量空间嵌入直线时的平均失真（即寻找维纳指数最小的生成路径） [11],，利用 [27], 可在欧几里得平面中得到一个 $(2+\varepsilon)$-近似解。

近期在几何背景下研究的一个相关问题是计算多边形的 Beer 指数 P。Beer 指数定义为两个随机（均匀）分布的点彼此可见的概率 P [1]；同一篇论文还研究了计算多边形中两个随机点之间期望距离的问题。与维纳指数一样，该问题基于计算所有点对之间距离之和（在连续情形下通过积分求值）。

另一个与维纳指数相关的研究领域是伸展子图：给定带权图 G 和实数 $t>1$， G 的一个 t-伸展子图是 G的生成子图 G^*，满足对于 G中的任意两个顶点 u和 v， $\delta_{G^*}(u,v) \le t\cdot\delta_G(u,v)$。因此， G^* 中的最短路径距离近似于底层图 G中的最短路径距离，而参数 t 表示近似比。使得 G^* 成为 G 的 t-伸展子图的最小 t 称为伸展因子。关于伸展子图的文献浩如烟海，尤其是在几何领域（例如参见 [4–7, 13, 22, 25]）。在几何图中， G中两个顶点 u 和 v之间的伸展因子，是 G 中从 u 到 v 的最短路径的欧几里得长度与 u 和 v之间欧几里得距离的比值。 G 的平均伸展因子是对 G中所有顶点对的伸展因子取平均值。对于一个给定的带权连通图 $G=(V,E)$，其边权均为正，且给定一个正值 W，平均伸展因子生成树问题旨在寻找 G 的一棵生成树 T，使得其平均伸展因子（对 $\binom{n}{2}$ 个顶点对取平均）不超过 W。对于欧几里得平面上的点，可以在多项式时间内构造一棵平均伸展因子为常数的生成树 [9]。

2 预备知识

设 P 是平面中的 n 个点的集合，并设 $G=(P,\ E)$ 是以 P 为顶点集的完全图。对于每条边 $(p,\ q)\in E$，令 $w(p,\ q)=|pq|$ 表示 $(p,\ q)$ 的权重，

其值为 p 与 q 之间的欧几里得距离 $|pq|$。设 T 是 P 的一棵生成树。对于点 p，$q \in P$，令 $\delta_T(p, q)$ 表示 p 与 q 之间在 T 中的（唯一）路径的权重。令 $W(T) = \sum_{p,\ q\in P} \delta_T(p, q)$ 表示 T 的维纳指数，即 T 中任意一对点之间路径的权重之和。最后，对于点 $p \in P$，令 $\delta_p(T) = \sum_{q\in P} \delta_T(p, q)$ 表示 T 中从 p 到 P 中每个点的路径权重总和。

定理 1. 设 T^* 是使维纳指数最小的 P 生成树。那么，T^* 是平面的。

证明。反设 T 中有两条边 (a, c) 和 (b, d) 相交。将边 (a, c) 和 (b, d) 从 T 中移除，得到森林 F。因此，F 包含三棵子树。不妨假设 a 和 b 位于同一棵子树 T_{ab} 中，而 c 和 d 分别位于不同的子树 T_c 和 T_d 中；见图 1。设 n_{ab}、n_c 和 n_d 分别为 T_{ab}、T_c 和 T_d 中点的数量。因此，

$$\begin{aligned} W(T^*) = {} & W(T_{ab}) + n_c \cdot \delta_a(T_{ab}) + n_d \cdot \delta_b(T_{ab}) \\ & + W(T_c) + (n_{ab} + n_d) \cdot \delta_c(T_c) + n_c(n_{ab} + n_d) \cdot |ac| \\ & + W(T_d) + (n_{ab} + n_c) \cdot \delta_d(T_d) + n_d(n_{ab} + n_c) \cdot |bd| \\ & + n_c \cdot n_d \cdot \delta_{T^*}(a, b)\,. \end{aligned}$$

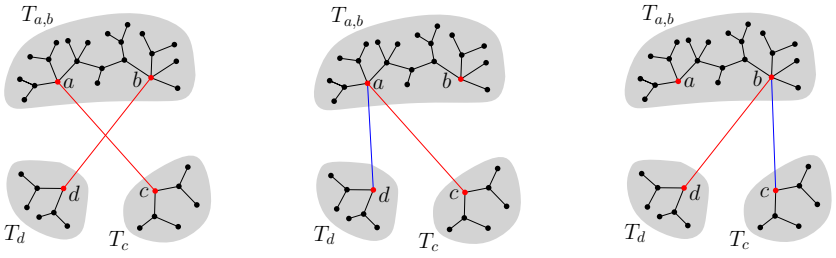


图 1. 树 T^*、T' 和 T''（从左到右）。

令 T' 为 P的一棵生成树，它由 T^* 通过将边 (b, d) 替换为边 (a, d) 得到。类似地，令 T'' 为 P的一棵生成树，它由 T^* 通过将边 (a, c) 替换为边 (b, c) 得到。因此，

$$\begin{aligned} W(T') = {} & W(T_{ab}) + (n_c + n_d) \cdot \delta_a(T_{ab}) \\ & + W(T_c) + (n_{ab} + n_d) \cdot \delta_c(T_c) + n_c(n_{ab} + n_d) \cdot |ac| \\ & + W(T_d) + (n_{ab} + n_c) \cdot \delta_d(T_d) + n_d(n_{ab} + n_c) \cdot |ad|\,, \end{aligned}$$

and

$$\begin{aligned} W(T'') = W(T_{ab}) &+ (n_c + n_d) \cdot \delta_b(T_{ab}) \\ &+ W(T_c) + (n_{ab} + n_d) \cdot \delta_c(T_c) + n_c(n_{ab} + n_d) \cdot |bc| \\ &+ W(T_d) + (n_{ab} + n_c) \cdot \delta_d(T_d) + n_d(n_{ab} + n_c) \cdot |bd| \,. \end{aligned}$$

因此，

$$\begin{aligned} W(T^*) - W(T') = n_d\big(\delta_b(T_{ab}) - \delta_a(T_{ab})\big) &+ n_d(n_{ab} + n_c)\big(|bd| - |ad|\big) \\ &+ n_c \cdot n_d \cdot \delta_{T^*}(a, b) \,, \end{aligned}$$

and

$$\begin{aligned} W(T^*) - W(T'') = n_c\big(\delta_a(T_{ab}) - \delta_b(T_{ab})\big) &+ n_c(n_{ab} + n_d)\big(|ac| - |bc|\big) \\ &+ n_c \cdot n_d \cdot \delta_{T^*}(a, b) \,. \end{aligned}$$

如果 $W(T^*) - W(T') > 0$或 $W(T^*) - W(T'') > 0$，那么这与 T^* 的极小性矛盾，结论得证。

假设 $W(T^*) - W(T') \le 0$且 $W(T^*) - W(T'') \le 0$。由于 $n_c > 0$和 $n_d > 0$，我们有

$$\delta_b(T_{ab}) - \delta_a(T_{ab}) + (n_{ab} + n_c)\big(|bd| - |ad|\big) + n_c \cdot \delta_{T^*}(a, b) \le 0 \,,$$

and

$$\delta_a(T_{ab}) - \delta_b(T_{ab}) + (n_{ab} + n_d)\big(|ac| - |bc|\big) + n_d \cdot \delta_{T^*}(a, b) \le 0 \,.$$

因此，将这些不等式相加可得

$$(n_{ab} + n_c)\big(|bd| - |ad|\big) + (n_{ab} + n_d)\big(|ac| - |bc|\big) + (n_c + n_d) \cdot \delta_{T^*}(a, b) \le 0 \,.$$

也就是说，

$$\begin{aligned} n_{ab}\big(|bd| + |ac| - |ad| - |bc|\big) + n_c\big(|bd| + \delta_{T^*}(a, b) - |ad|\big) \\ + n_d\big(|ac| + \delta_{T^*}(a, b) - |bc|\big) \le 0 \,. \end{aligned}$$

由于 $n_{ab}, n_c, n_d > 0$，并且根据三角不等式，$|bd| + |ac| - |ad| - |bc| > 0$,$|bd| + \delta_{T^*}(a, b) - |ad| > 0$，以及 $|ac| + \delta_{T^*}(a, b) - |bc| > 0$，这就产生了矛盾。

□

3 凸位置点的精确算法

令 $\{p_1, p_2, \ldots, p_n\}$ 表示按顺时针顺序连接 P 中的点所得到的凸多边形的顶点，其中任意选取 p_1 作为第一个点；参见图 2。为简洁起见，我们假设所有下标均模 n 计算。对于每个 $1 \le i \le j \le n$，令 $P[i, j] \subseteq P$ 为集合 $\{p_i, p_{i+1}, \ldots, p_j\}$。令 $T_{i,j}$ 为 $P[i, j]$ 的一棵生成树，并令 $W(T_{i,j})$ 表示

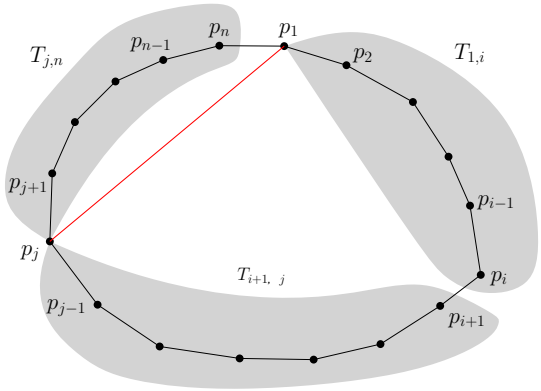


图2。 由 P 得到的凸多边形。 p_1 在 T^* 中与 p_j 相连。

其维纳指数。对于一个点 $x \in \{i, j\}$，令 $\delta_x(T_{i,j})$ 表示在 $T_{i,j}$ 中，从 p_x 到 $P[i, j]$ 中每个点的最短路径的总权重。也就是说，$\delta_x(T_{i,j}) = \sum_{p \in P[i,j]} \delta_{T_{i,j}}(p_x, p)$。

设 T^* 是 P 的最小维纳指数树，其维纳指数为 W^*。注意，对于任意 $1 \le i < j \le n$，$P[i, j]$ 中的点都处于凸位置，因为 P 中的点都处于凸位置。由于 T^* 是一棵生成树，每个点，尤其是 p_1，都至少与 T^* 中的一条边相邻。设 p_j 是索引最大的点，其索引为 j，并且它在 T^* 中与 p_1 相连。此外，存在一个索引 $1 \le i \le j$，使得 $P[1, i]$ 中的所有点在 T^* 中都比起 p_j 更接近 p_1，而 $P[i + 1, j]$ 中的所有点在 T^* 中都比起 p_1 更接近 p_j。因此，

$$
\begin{aligned}
W^* = {} & W(T_{1,i}) + (n-i) \cdot \delta_1(T_{1,i}) && (1)\\
& + W(T_{i+1,j}) + (n-j+i) \cdot \delta_j(T_{i+1,j}) && (2)\\
& + W(T_{j,n}) + (j-1) \cdot \delta_j(T_{j,n}) && (3)\\
& + i(n-i) \cdot |p_1 p_j|. && (4)
\end{aligned}
$$

因此，为了计算 W^*，我们对每个介于 2 和 n 之间的 i 以及每个介于 1 和 i 之间的 j，分别计算 (1)、(2)、(3) 和 (4)，并取这些值之和的最小值。一般而言，对于每个 $1 \le i < j \le n$，令 $W_j[i, j] = W(T_{i,j}) + (n-j+i-1) \cdot \delta_j(T_{i,j})$ 表示以 p_j 为根的 $P[i, j]$ 的生成树 $T_{i,j}$ 所能取得的最小值。同样地，令 $W_i[i, j] = W(T_{i,j}) + (n-j+i-1) \cdot \delta_i(T_{i,j})$ 表示以 p_i 为根的 $P[i, j]$ 的生成树 $T_{i,j}$ 所能取得的最小值。因此，我们可以使用以下公式递归地计算 $W_j[i, j]$ 和 $W_i[i, j]$；另请参见图 3。

$$
W_j[i,j] = \min_{\substack{i \le k < j \\ k \le l < j}} \left\{ W_k[i,k] + W_k[k,l] + W_j[l+1,j] + (l-i+1)(n-l+i-1) \cdot |p_k p_j| \right\},
$$

and

$$W_i[i,j] = \min_{\substack{i<k\le j \\ i\le l<k}} \left\{ W_i[i,l] + W_k[l+1,k] + W_j[k,j] + (j-l)(n-j+l) \cdot |p_i p_k| \right\}.$$

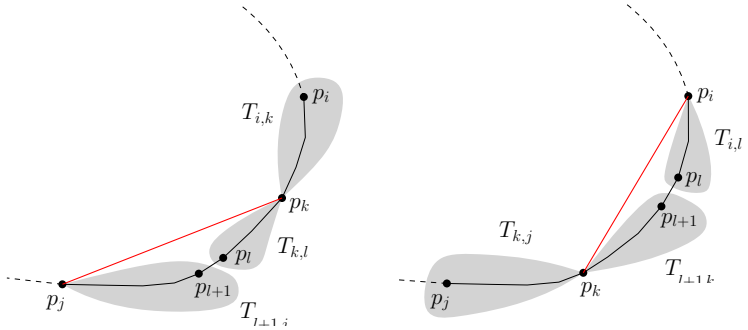


图3. 由 $P[i,j]$定义的子问题。(a) 计算 $W_j[i,j]$。(b) 计算$W_i[i,j]$。

我们使用动态→ ←规划如下计算每个 $1 \le i < j \le n$ 对应的 $W_j[i,j]$ 和 $W_i[i,j]$。我们维护两个表 M 和 M，每个表的大小均为 $n \times n$，使得对于每个 $1 \le i < j \le n$，都有 $\overrightarrow{M}[i,j] = W_j[i,j]$ 和 $\overleftarrow{M}[i,j] = W_i[i,j]$。我们使用算法 **1** 填充这些表。

算法 1 *ComputeOptimal*(P)

1: $n \leftarrow |P|$
2: **对于**每个 $i \leftarrow 1$ 到 n **执行**
$\overrightarrow{M}[i,i] \leftarrow 0$
$\overleftarrow{M}[i,i] \leftarrow 0$
3: **对于**每个 $j \leftarrow n$ 到 **1 执行**
对于每个 $i \leftarrow j$ 到 n **执行**

$$\overrightarrow{M}[i,j] \leftarrow \min_{\substack{i\le k<j \\ k\le l<j}} \left\{ \overrightarrow{M}[i,k] + \overleftarrow{M}[k,l] + \overrightarrow{M}[l+1,j] + (l-i+1)(n-l+i-1) \cdot |p_k p_j| \right\}$$

$$\overleftarrow{M}[i,j] \leftarrow \min_{\substack{i<k\le j \\ i\le l<k}} \left\{ \overleftarrow{M}[i,l] + \overrightarrow{M}[l+1,k] + \overrightarrow{M}[k,j] + (j-l)(n-j+l) \cdot |p_i p_k| \right\}$$

4: **返回** $\overleftarrow{M}[1,n]$

→ → ←请注意，当我们填入单元格 $M[i,j]$时，对于每个 $i \le k < j$ 和每个 $k \le l < j$，所有单元格 $M[i,k]$、 $M[k,l]$和 $\rightarrow M[l+1,j]$都已计算完毕，

而当我们填入单元格 $\overleftarrow{M}[i,j]$ 时，对于每个 $i<k\le j$ 和每个 $i\le l<k$，所有单元格 $\overleftarrow{M}[i,l]$、$\overrightarrow{M}[l+1,k]$ 和 $\overrightarrow{M}[k,j]$ 都已计算完毕。因此，表中的每个单元格都能在 $O(n^2)$ 时间内计算完成，而整张表则能在$O(n^4)$ 时间内计算完成。

以下定理概述了本节的结果。

定理 2. 设 P 是一组处于凸位置的 n 点。那么，可以在 $O(n^4)$ 时间内计算出 P 的最小维纳指数生成树。

4 困难性证明

设 P 是平面上的点集，T 是 P 的一棵生成树。我们将 T 的维纳指数定义为 $W(T)=\sum_{p,\ q\in P}\delta_T(p,q)$，并将 T 的权重定义为 $wt(T)=\sum_{(p,q)\in T}|pq|$，其中 $\delta_T(p,q)$ 是 T 中 p 与 q 之间路径的长度，$|pq|$ 是 p 与 q 之间的欧几里得距离。对于一条边 (p,q)，令 $N_T(p)$（分别为 $N_T(q)$）表示 T 中距离 q 比距离 q 更近（分别是距离 q 比距离 p 更近）的点的数量。众所周知，$W(T)$ 可以表示为:

$$W(T)=\sum_{(p,q)\in T}N_T(p)\cdot N_T(q)\cdot|pq|.$$

在本节中，我们证明以下问题是 NP 困难的。

欧几里得维纳指数树问题：给定平面上的一组点 P、一个成本 W 和一个预算 B，判断是否存在 P 的生成树 T，使得 $W(T)\le W$ 且 $wt(T)\le B$。

定理 3。欧几里得维纳指数树问题是弱 *NP* 困难的。

证明。受 Carmi 和 Chaitman-Yerushalmi [8], 的启发，我们将已知为 NP 困难的划分问题 [14], 归约到欧几里得维纳指数树问题。在划分问题中，给定一组 $X=\{x_1,x_2,\ldots,x_n\}$ 个 n正整数，其总和为偶数 $R=\sum_{i=1}^{n}x_i$，目标是判断是否存在一个子集 $S\subseteq X$，使得 $\sum_{x_i\in S}x_i=\frac{1}{2}R$。

给定一个划分问题的实例 $X=\{x_1,x_2,\ldots,x_n\}$，其中 x_i 是整数，我们按如下方式构造一个包含 $m=n^3+3n$ 个点的集合 P。集合 P 由 n 个点组成：p_1，p_2、.....、p_n，这些点在半径为 nR 的圆上等距分布；另有一个包含 n^3 个点的簇 C，位于圆心。此外，对每个$1\le i\le n$，我们放置两个点 l_i 和 r_i，它们到 p_i 的距离均为 x_i，且彼此间的距离为 $\frac{1}{2}x_i$；见图 4。最后，我们设定

$$\begin{aligned}B&=\Big(n^2+\frac{7}{4}\Big)R,\text{ and}\\W&=3n^2\big(m-3\big)R+\Big(\frac{9}{4}m-\frac{13}{4}\Big)R\\&=3n^5R+\frac{45}{4}n^3R-9n^2R+\frac{27}{4}nR-\frac{13}{4}R.\end{aligned}$$

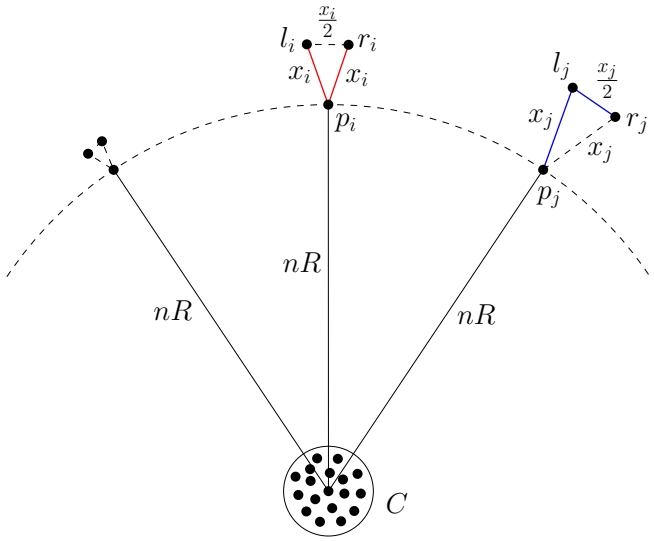


图 4. 归约生成的集合 P。将点 l_j、r_j 和 p_j 相连，得到 $x_j \in S$（蓝色）；将点 l_i、r_i 和 p_i 相连，得到 $x_i \in X \setminus S$（红色）。

假设存在一个集合 $S \subseteq X$，使得 $\sum_{x_i \in S} x_i = \frac{1}{2}R$。我们按如下方式为 P 中的点构造一棵生成树 T：

– 任取一个点 $s \in C$，并将其与 $C \cup \{p_1, p_2, \ldots, p_n\}$ 中的所有点相连，构成一棵以 s为中心的星形树。– 对于每个 $1 \le i \le n$，连接点 p_i 和 l_i。– 对于每个 $x_i \in S$，连接点 p_i 和 r_i。– 对于每个 $x_i \in X \setminus S$，连接点 r_i 和 l_i；见图 4。

很容易看出，$wt(T) = n^2R + R + \frac{3}{4}R = \left(n^2 + \frac{7}{4}\right)R = B$。此外，$T$ 的维纳指数为：

$$\begin{aligned}
W(T) &= \sum_{(p,q)\in T} N_T(p) \cdot N_T(q) \cdot |pq| \\
&= 3(n^3+3n-3)n^2R + \sum_{x_i \in S'} 2(n^3+3n-1)x_i \\
&\quad + \sum_{x_i \notin S'} \left((n^3+3n-1)\frac{1}{2}x_i\right) + \sum_{x_i \notin S'} \left(2(n^3+3n-2)x_i\right) \\
&= 3n^5R + 9n^3R - 9n^2R + (n^3+3n-1)R \\
&\quad + \frac{1}{4}(n^3+3n-1)R + (n^3+3n-2)R \\
&= 3n^5R + \frac{45}{4}n^3R - 9n^2R + \frac{27}{4}nR - \frac{13}{4}R = W\,.
\end{aligned}$$

反之，设 T' 是 P的生成树，其中 $wt(T') \le B$，且 $W(T') \le W$。

断言。 满足 $p \in C$和 $q \in P \setminus C$的边 $(p,q) \in T'$的数量为 n。

证明。 假设有 k 条这样的边。每条这样的边的权重至少为nR，因此 $wt(T') \geq knR$，由于 $B = (n^2 + \frac{7}{4})R$，我们得到 $k \leq n$。我们有

$$\begin{aligned} W(T') &> (3knR + 3(n-k)(nR + 2\pi R))n^3 \\ &= (3kn + 3n^2 + 6n\pi - 3kn - 6k\pi)n^3 R \\ &= (3n^2 + 6\pi(n-k))n^3 R \\ &= 3n^5 R + 6\pi(n-k)n^3 R\,. \end{aligned}$$

因此，如果 $k < n$，那么我们得到 $W(T') > 3n^5 R + 6\pi n^3 R > W$，其中 n充分大。□

令 $G_i = \{p_i, l_i, r_i\}$，对每个 $1 \leq i \leq n$ 均成立。根据断言 4 的证明可知，对每个 $1 \leq i \leq n$，在 T' 中恰好存在一条边（p，q），其中$q \in G_i$ 且 $p \in C$。此外，很容易看出 $q = p_i$。因此，在每个 G_i 中，我们都有（p_i，l_i）$\in T'$或（p_i，r_i）$\in T'$。不失一般性，假设（p_i，l_i）$\in T'$。因此，（p_i，r_i）$\in T'$或（l_i，r_i）$\in T'$。令 $S' \subseteq X$，使得当且仅当（p_i，r_i）$\in T'$时，$x_i \in S'$，并令 $R' = \sum_{x_i \in S'} x_i$。

因此，为完成证明，我们证明：如果 $R' \neq \frac{1}{2}R$，那么要么 $wt(T') > B$，要么 $W(T) > W$。

情况 1： $R' > \frac{1}{2}R$。在这种情况下，我们有

$$\begin{aligned} wt(T') &\geq n^2 R + \sum_{x_i \in S'} 2x_i + \sum_{x_i \notin S'} \frac{3}{2}x_i \;=\; n^2 R + 2R' + \frac{3}{2}(R - R') \\ &= n^2 R + \frac{1}{2}R' + \frac{3}{2}R \;>\; n^2 R + \frac{1}{4}R + \frac{3}{2}R \;=\; \big(n^2 + \frac{7}{4}\big)R \;= B\,. \end{aligned}$$

因此，$wt(T') > B$。

情况 2： $R' < \frac{1}{2}R$。在这种情况下，我们有

$$\begin{aligned} W(T) &= \sum_{(p,q) \in T} N_T(p) \cdot N_T(q) \cdot |pq| \\ &= 3(n^3 + 3n - 3)n^2 R + \sum_{x_i \in S'} 2(n^3 + 3n - 1)x_i \\ &\quad + \sum_{x_i \notin S'} \Big((n^3 + 3n - 1)\frac{1}{2}x_i\Big) + \sum_{x_i \notin S'} \Big(2(n^3 + 3n - 2)x_i\Big) \\ &= 3n^5 R + 9n^3 R - 9n^2 R + 2(n^3 + 3n - 1)R' \\ &\quad + \frac{1}{2}\Big(n^3 + 3n - 1\Big)(R - R') + 2(n^3 + 3n - 2)(R - R') \end{aligned}$$

$$
\begin{aligned}
&= 3n^5R + 9n^3R - 9n^2R + 2(n^3+3n-2)R \\
&\quad - \Big(\frac{1}{2}\Big(n^3+3n-1\Big) - 2\Big)R' + \frac{1}{2}\Big(n^3+3n-1\Big)R \\
&\quad - \Big(\frac{1}{2}\Big(n^3+3n-1\Big) - 2\Big)R' + \frac{1}{2}\Big(n^3+3n-1\Big)R \\
&> 3n^5R + 9n^3R - 9n^2R + 2(n^3+3n-2)R \\
&\quad - \frac{1}{2}\Big(\frac{1}{2}\Big(n^3+3n-1\Big) - 2\Big)R + \frac{1}{2}\Big(n^3+3n-1\Big)R \\
&= 3n^5R + \frac{45}{4}n^3R - 9n^2R + \frac{27}{4}nR - \frac{13}{4}R = W\,.
\end{aligned}
$$

5 条优化维纳指数的路径

现在考虑使维纳指数最优的生成路径的情况。

定理 4. 设 P 是由 n 个点组成的集合。在 P 的所有哈密顿路径中，使维纳指数最小的路径不一定是平面的。

证明。 考虑图 5 所示的凸位置点集 P，其中包含 $n = 2m + 2$ 个点。集合 P 由两个簇 P_l 和 P_r 以及两个点 p 和 q组成，其中 $|P_l| = |P_r| = m$。簇 P_l 中的点任意接近原点 (0, 0)，簇 P_r中的点任意接近坐标 (6,0)。点 p 位于坐标 (5,1)，点 q 位于坐标 $(5, -1)$。



图 5. 凸位置的 $n = 2m + 2$ 个点组成的集合 P。

为简化计算，我们假设连接 P_l 中各点的路径的维纳指数为零，连接 P_r 中各点的路径的维纳指数也为零。因此，任何旨在使维纳指数最小的 P 点路径 Π 都会用一条路径连接 P_l 中的点，并用一条路径连接 P_r 中的点。我们计算了定义在点 (0,0)、(6, 0)、 p和 q上的所有可能哈密顿路径的维纳指数；见图 6。计算结果表明，维纳指数最小的哈密顿路径不是平面的（当 n足够大时）。□

定理 5。 对于欧几里得平面中的点，计算维纳指数最小的哈密顿路径是 *NP* 困难的。

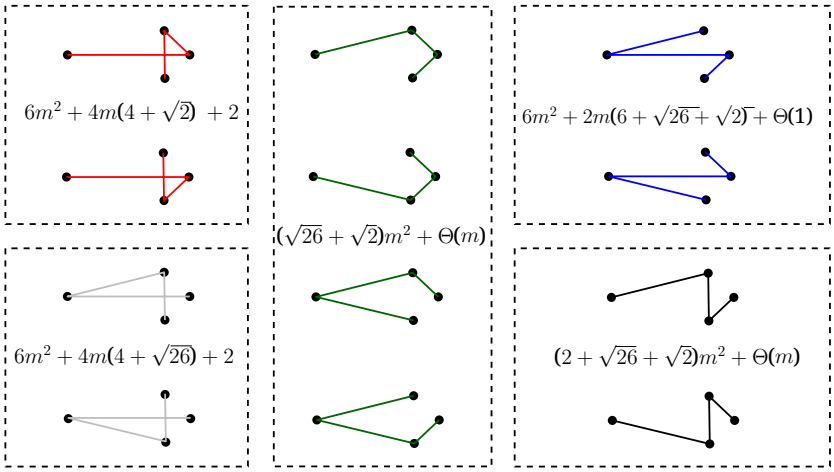


图 6. 在点 (0，0)、(6，0)、 p 和 q 上定义的 12 条可能的哈密顿路径的维纳指数（假设位于 (0，0) 上的 m 点由一条路径连接，位于 (6，0) 上的 m 点由一条路径连接，且两者的维纳指数均为零）。

证明。 我们从网格图中的哈密顿性问题归约而来（网格图的顶点是整数网格点，边连接距离为一的网格点对）。首先注意，对于 n 个点组成且每条边长度均为一的哈密顿路径，其维纳指数为 $\sum_{i=1}^{n-1} i(n-i) = \binom{n+1}{3}$；见图 7。因此，很容易看出，网格图 G 存在哈密顿路径，当且仅当存在一条维纳指数为 $\binom{n+1}{3}$的路径。□

定理 6. 存在一个平面点集 P，使得任意哈密顿路径的维纳指数至少为完全欧几里得图 P 的维纳指数的 $\Theta(\sqrt{n})$ 倍。

证明。 设 P是位于$\sqrt{n} \times \sqrt{n}$整数网格上的一组 n点。P 的任意哈密顿路径的维纳指数至少为 $\binom{n+1}{3}$，这也是所有边长度均为 1 的哈密顿路径的维纳指数。因此，P 的任意哈密顿路径的维纳指数至少为 $\Theta(n^3)$。另一方面， P 上的完全图的维纳指数为 $\Theta(n^{2.5})$。□

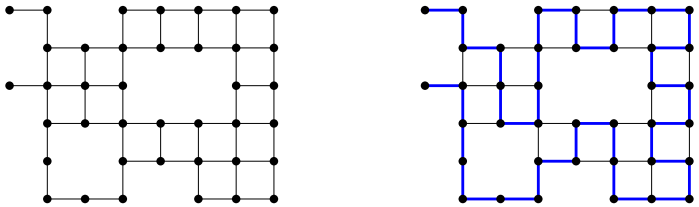


图 7. 一个网格图 G 以及其中一条维纳指数为 $\binom{n+1}{3}$ 的哈密顿路径。

参考文献

1. M. Abrahamsen 和 V. Fredslund-Hansen. 多边形的凸度与期望距离。*arXiv* 预印本 *arXiv:2208.07106*，2022年。2. D. Bonchev. 维纳数：若干应用与最新进展。载于化学中的拓扑学，第 58–88 页，2002年。3. D. Bonchev 和 N. Trinajstić. 信息论、距离矩阵与分子分支。《化学物理学杂志》，67(10):4517–4533，1977年。4. P. Bose、P. Carmi 和 L. Chaitman-Yerushalmi. 关于有界度平面强几何伸展子。《离散算法杂志》，15:16–31，2012年。5. P. Bose、J. Gudmundsson 和 M. H. M. Smid. 构造有界度、低权重的平面伸展子。载于*ESA*，第 234–246 页，2002年。6. P. Bose、D. Hill 和 M. H. M. Smid. 低度平面伸展子的改进伸展比。《算法》，80(3):935–976，2018年。7. J. Cardinal、S. Collette 和 S. Langerman. 几何图的局部性质。《计算几何》，39(1):55–64，2008年。8. P. Carmi 和 L. Chaitman-Yerushalmi. 最小权重欧几里得 t-伸展子是 NP 困难的。《离散算法杂志》，22:30–42，2013年。9. S.-W. Cheng、C. Knauer、S. Langerman 和 M. H. M. Smid. 近似计算几何图的平均伸展因子。《计算几何杂志》，3(1):132–153，2012年。10. K. Das、I. Gutman 和 B. Furtula. 图的几何算术指标综述。《*Match*（德国米尔海姆）》，65:595–644，2011年1月。11. K. Dhamdhere、A. Gupta 和 R. Ravi. 最小化平均失真的近似算法。《计算系统理论》，39(1):93–111，2006年。12. A. A. Dobrynin、R. Entringer 和 I. Gutman. 树的维纳指数：理论与应用。《应用数学学报》，66(3):211–249，2001年。13. A. Filtser 和 S. Solomon. 贪心伸展子在存在性意义下是最优的。载于*2016* 年 *ACM* 分布式计算原理研讨会（*PODC*）论文集，第 9–17 页，2016年。14. M. R. Garey 和 D. S. Johnson. 《计算机与难解问题：*NP* 完全性理论指南》。W. H. Freeman & Co.，美国，1979年。15. T. F. Gonzalez. 《近似算法与元启发式手册》第 *59* 章。Chapman & Hall/CRC，2007年。16. A. Graovac 和 T. Pisanski. 关于图的维纳指数。《数学化学杂志》，8:53–62，1991年。17. F. Harary. 图论。Addison-Wesley，1969年。18. T. C. Hu. 最优通信生成树。《*SIAM* 计算杂志》，3(3):188–195，1974年。19. D. S. Johnson、J. K. Lenstra 和 A. H. G. Rinnooy Kan. 网络设计问题的复杂性。《网络》，8:279–285，1978年。20. L. Kier. 化学与药物研究中的分子连接性。爱思唯尔，1976年。21. M. Knor、R. Škrekovski 和 A. Tepeh. 维纳指数的数学方面。《当代数学艺术》，11:327–352，2015年。22. X. Y. Li 和 Y. Wang. 高效构造低权重有界度平面伸展子。《国际计算几何与应用杂志》，14(1-2):69–84，2004年。23. O. Mekenyan、D. Bonchev 和 N. Trinajstić. 具有非环分支的环状系统的结构复杂度与分子性质。《克罗地亚化学学报》，56(2):237–261，1983年。24. B. Mohar 和 T. Pisanski. 如何计算图的维纳指数。《数学化学杂志》，2(3):267，1988年。

14 A.K. Abu-Affash、P.Carmi、O.Luwisch 和 J. S.B. Mitchell

25. G. Narasimhan 和 M. Smid。《几何跨距网络》。剑桥大学出版社，2007年。26. S. Ronghua。树的平均距离。《系统科学与复杂性杂志》，6:18–24，1993年。27. R. Sitters。旅行维修工问题及其他最小延迟问题的多项式时间近似方案。《*SIAM* 计算杂志》，50 1602，2021年 5 1580( ): – 。28. N. Trinajstić。《化学中的数学与计算概念》。Ellis Horwood，奇切斯特，1986年。29. E. Weiszfeld 和 F. Plastria。关于使到给定 n 个点的距离之和最小的点。《运筹学年刊》，167(1):7–41，2009年。30. H. Wiener。石蜡沸点的结构测定。《美国化学会志》，69(1):17–20，1947年。31. B. Ye Wu、G. Lancia、V. Bafna、K.-M. Chao、R. Ravi 和 C. Y. Tang。最小路由成本生成树的多项式时间近似方案。《*SIAM* 计算杂志》，29(3):761–778，2000年。32. K. Xu、M. Liu、K. Das、I. Gutman 和 B. Furtula。关于距离型拓扑指数的极值图综述。《*Match*（德国米尔海姆）》，71:461–508，2014年 02月。