On the Minimum Spanning Tree Determined by n Points in the Unit Square

Ye Jichang(叶继昌) Xu Yinfeng(徐寅峰) Xu Chengxian(徐成贤)
(Xi'an Jiaotong University, Xi'an, 710049)

Abstract Let P_n be a set of n points in the unit square S, $l(P_n)$ denoe the length of the minimum spanning tree of P_n, and

$$C_n = \max_{P_n \subseteq S} l(P_n), \qquad n = 2, 3, \cdots$$

In this paper, the exact value of C_n for $n = 2, 3, 4$ and the corresponding configurations are given. Additionally, the conjectures of the configuration for $n = 5, 6, 7, 8, 9$ are proposed.

Keywords minimum spanning tree; maximin problem; configuration

§1. Introduction

A minimum spanning tree (MST) is widely applied to the fields of computer, communication, nerwork and so on. Many results have been obtaines, but few of them deal with the worst-case analysis for the given finite region. In fact, it is a maximin problem (see [1]——[3]). This paper is devoted to the worst-case of MST in the unit square.

Let S denote the unit square, P_n the set of n point in S, $l(P_n)$ be the length of a MST of P_n. The distance between two points is of Euclidean sense. Our problem is to determine C_n defined as follows

$$C_n = \max_{P_n \subseteq S} l(P_n), \qquad n = 2, 3, \cdots.$$

and the point set location of P_n^* for which

$$C_n = l(P_n^*), \quad P_n^* \subseteq S$$

is called the configuration of C_n. Let T_n demote the set of all possible spanning trees with vertices P_n and the length of t for $t \in T_n$, $l(t)$, we have

$$C_n = \max_{P_n \subseteq S} \min_{t \in T_n} l(t)$$

T_n^* is called an optimal tree if the length of T_n^* equals C_n.

This paper is organized as follows: in Section 2, the configration of C_n in the general case is discussed. The special cases for $n = 2, 3, 4$ are investigated in Section 3. Finally, the conjectures for $n = 5, 6, 7, 8, 9$ are proposed in Section 4.

Received Sept. 8, 1998.

§2 General Case

Let $P_n^* \subseteq S$, $C_n = l(P_n^*)$, $CH(P_n^*)$ be the convex hull of P_n^* and $V(P_n^*) = P_n^* \cap CH(P_n^*)$.

Theorem 1 On every edge of S there must be at least one point of P_n^*.

Proof As shown in Fig. 1(a), suppose that no points in P_n^* is on AD and p^* in P_n^* is the nearest point from AD. Let straight line l_{p^*} be parallel to DC and $\overline{P}^*$ be the crossing point of AD and l_{p^*}. If we only move p^* to $\overline{P}^*$, it is obvious that

$$l(P_n^*) < l(P_n^* \cup \{\overline{p}^*\} \setminus \{p^*\}).$$

This is a contradiction with the definition of P_n^*.

(a) (b)

Fig. 1

Theorem 2 Points in $V(P_n^*)$ must be on the boundary of S.

Proof Suppose that $p^* \in V(P_n^*)$ and p^* is an interior point of S. By Theorem 1, let p_i and p_j be the two points in $V(P_n^*)$ that connect with p^* in CH(P_n^*) and on the boundary of S. Let straight line l_{p^*} be the bisector of angle $p_i p^* p_j$. If we only move p^* in direction as shown in Fig. (b). Write $\overline{p}^*$ the crossing point that l_{p^*} and S, we have

$$l(P_n^*) < l(P_n^* \cup \overline{p}^* \setminus P^*),$$

which is a contradiction with the definition of P_n^*.

Theorem 3 $C_{n+1} \geq C_n$.

It is an obvioud conclusion, so the proof is omitted.

Theorem 4 $\lim_{n\to\infty} C_n = \infty$.

Proof Divide the unit square into $(n-1)^2$ small equal squares with edge lengh $1/(n-1)$. The total number of vertices of all small square is n^2. Consider the MST with the n^2 vetices P_{n^2}, it is obviius that $l(P_{n^2}^*)$ is $n+1$. Therefore

$$l(P_{n^2}^*) \geq l(P_{n^2}) = n + 1.$$

It follows from Theorem 3 that $\lim_{n\to\infty} C_n = \infty$.

§3 Special Cases

The objective of this section is to determine the lication of P_n^* for $n = 2, 3, 4$.

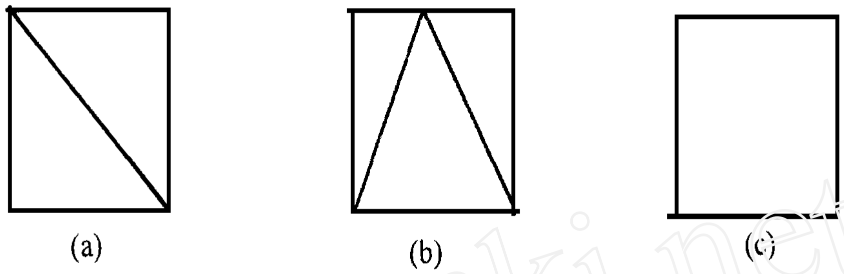

Fig 2

Theorem 5 For $n=2,3,4$ the location of P_n^* is shown in Fig 2

Proof For $n=2$ the location is trivial and $C_2=\sqrt{2}$. It is sufficiency only to prove $n=3$ and $n=4$

First consider the case $n=3$, we shall prove that the location P_3^* as shown in Fig 2(b) is optimal and $C_3=1+\frac{\sqrt{5}}{2}$.

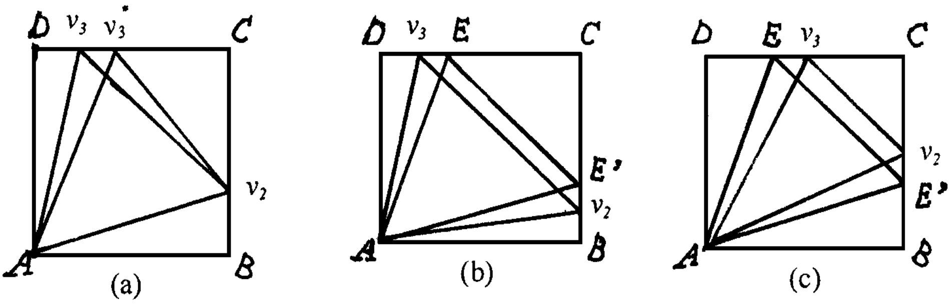

Fig 3

Form Theorem 1 and Theorem 2 we know that CH (P_3^*) is a triangle and there must be one point located at a corner of the square Let points A, v_2, v_3 be in P_3^*, as shown in Fig 3 (a). We claim that triangle Av_2v_3 is an isosceles triangle Conversely suppose that triangle Av_2v_3 is not an isosceles triangle If v_2v_3 is the longest edge, i e $|Av_3|<|v_2v_3|$ and $|v_2A|<|v_2v_3|$ ($|AB|$ denotes the Euclidean distance between A and B), then $l(P_3^*)=|Av_3|+|Av_2|$. Moving the point v_3 on the edge DC to v_2^* slightly enough we get $\overrightarrow{P_3}=\{A, v_2, v_3^*\}$ such that $|Av_3^*|\le|v_3^*v_2|$ and $|v_2A|\le|v_3^*v_2|$. So we have $l(\overrightarrow{P_3})>l(P_3^*)$. This is a contradiction with the definition of P_3^*. Similar we can give a contradiction that either Av_2 or Av_3 is the longest one So the triangle Av_2v_3 is an isosceles triangle

Suppose $|Av_3|=|Av_2|$, we can prove that triangle Av_2v_3 is an equilateral triangle and $l(P_3^*)=2(\sqrt{6}-\sqrt{2})$.

If point v_3 is in DE, as shown in Fig 3(b), it is obvious that

$$|v_2v_3|>|Av_3|=|Av_2|.$$

Hence, we have

$$|Av_3| + |Av_2| < |AE| + |AE'| = 2(\sqrt{6} - \sqrt{2}).$$

Let $|DE| = |BE'| = 2 - \sqrt{3}$, then $|AE| = |AE'| + |EE'| = \sqrt{6} - \sqrt{2}$. If point v_3 is in EC, as shown in Fig. 3(c), then

$$|v_2v_3| < |Av_3| = |Av_2|, \quad |Dv_3| = |Bv_2| > |DE| = |BE'| = 2 - \sqrt{3},$$
$$|Av_3| = (1 + (|DE| + |Ev_3|)^2)^{1/2}, \quad |AE| = (1 + |DE|^2)^{1/2},$$
$$|AE| + |EE'| - |Av_3| - |v_2v_3| = (|EE'| - |v_2v_3|) - (|Av_3| - |AE|)$$
$$= \sqrt{2}\,|Ev_3| - \frac{2|DE| \cdot |Ev_3| + |Ev_3|^2}{\sqrt{1 + |DE|^2} + \sqrt{1 + (|DE| + |Ev_3|)^2}}$$
$$> \sqrt{2}\,|Ev_3| - \frac{|Ev_3|(2|DE| + |Ev_3|)}{2\sqrt{1 + |DE|^2}}$$
$$> |Ev_3|(\sqrt{2} - \frac{1}{2}(2|DE| + |Ev_3|))$$
$$= |Ev_3|(\sqrt{2} - \frac{1}{2}(2 - \sqrt{3} + |Dv_3|))$$
$$\geq |Ev_3|(\sqrt{2} - \frac{3 - \sqrt{3}}{2})$$
$$> |Ev_3|(\sqrt{2} - 1) > 0.$$

So, we have $|Av_3| + |v_2v_3| < |AE| + |EE'| = 2(\sqrt{6} - \sqrt{2})$.

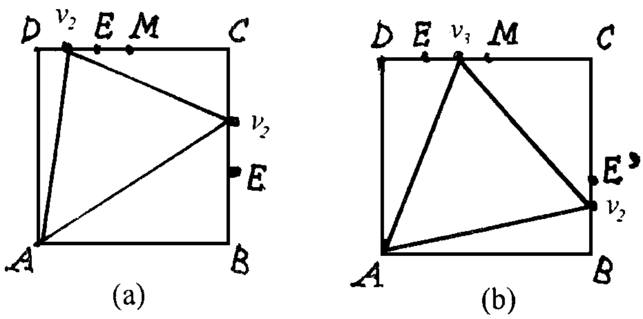

Fig. 4

Suppose that $|Av_3| = |v_2v_3|$. As shown in Fig. 4, the definition of point E is as above and point M is the midpoint of DC. If point v_2 is on DE, we have

$$|Av_3| = |v_2v_3| < |Av_2|, \; |Av_3| < |AE|, \; |Av_2| + |v_2v_3| < |AE| + |AE'|.$$

If point v_3 is on EM, as shown in Fig. 4(b), then $|Av_3| = |v_3v_2| > |Av_2|$, we shall show that $|Av_3| + |Av_2| < |AM| + |AB| = 1 + \frac{\sqrt{5}}{2}$.

Let $|v_3M| = y$, we have

$$0 < y < \sqrt{3} - \frac{3}{2}, \; |Dv_3| = \frac{1}{2} - y, \; |v_3C| = \frac{1}{2} + y,$$
$$|Av_3| = \sqrt{1 + |Dv_3|^2} = \sqrt{1 + (\frac{1}{2} - y)^2} = \frac{1}{2}\sqrt{5 - 4y + 4y^2},$$
$$|Cv_2| = \sqrt{|Av_3|^2 - |v_3C|^2} = \sqrt{1 + (\frac{1}{2} - y)^2 - (\frac{1}{2} + y)^2} = \sqrt{1 - 2y},$$

$$|v_2B| = 1 - |Cv_2| = 1 - \sqrt{1-2y},$$

$$|Av_2| = \sqrt{1 + |Bv_2|^2} = \sqrt{1 + (1 - \sqrt{1-2y})^2} = \sqrt{3 - 2y - 2\sqrt{1-2y}}.$$

Noticing the inequality

$$\sqrt{1-2y} > 1 - y - 2y^2, \quad 0 < y < \sqrt{3} - \frac{3}{2},$$

we get

$$|AM| - |Av_3| = \frac{\sqrt{5}}{2} - \frac{\sqrt{5 - 4y + 4y^2}}{2} = \frac{4y(1-y)}{2(\sqrt{5} + \sqrt{5 - 4y + 4y^2})} > \frac{y(1-y)}{\sqrt{5}},$$

$$\begin{aligned} |Av_2| - |AB| &= \sqrt{3 - 2y - 2\sqrt{1-2y}} - 1 \\ &= \frac{2 - 2y - 2\sqrt{1-2y}}{1 + \sqrt{3 - 2y - 2\sqrt{1-y}}} \\ &< \frac{2 - 2y - 2\sqrt{1-2y}}{2} \\ &= 1 - y - \sqrt{1-2y} < 1 - y - (1 - y - 2y^2) \\ &= 2y^2 < \frac{y(1-y)}{\sqrt{5}}. \end{aligned}$$

So, $|AM| - |Av_3| > |Av_2| - |AB|$, i e $|Av_2| + |Av_3| < |AM| + |AB| = 1 + \frac{\sqrt{5}}{2}$.

Observing that $2(\sqrt{6} - \sqrt{2}) < 1 + \frac{\sqrt{5}}{2})$ is follows from above discussion that the location of P_3^* is as shown in Fig 2(b) and $C_3 = 1 + \frac{\sqrt{5}}{2}$.

For $n = 4$ we shall prove that the location of P_4^* is as shown in Fig 2(c) and $C_4 = 3$.

It's known from Theorem 1 that there are at least two points on the boundary of the square.

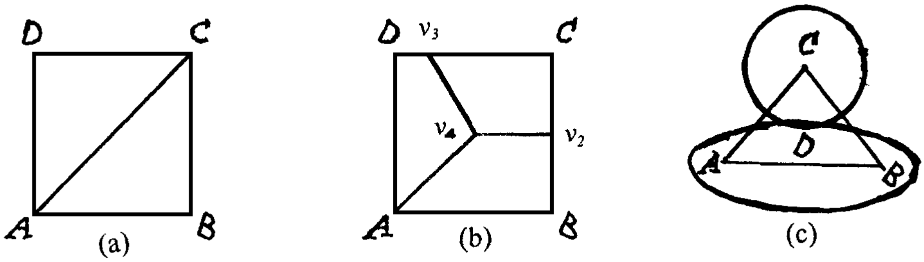

Fig 5

First we suppose that there are two points of P_4^* on the boundary. By Theorem 1 and 2 we know that the two points must coincide with A and C as shown Fig 5(a) and the other two points of P_4^* are on the line segment AC. Hence $l(P_4^*) < 3$.

Secondly, suppose that there are three points on the boundary of S. From Theorem 1 and 2, we know that the hull of P_4^* is a triangle and there is one point of P_4^* in the interior of the triangle as Fig. 5(b).

Lemma Given any triangle ABC, such that $|AB| \geq |BC| \geq |AC|$ and D is in the interior of the triangle ABC. Then the following inequality holds

$$|AD| + |BD| + |CD| \leq 2|AB|.$$

Proof As shown Fig. 5(c) we construct an ellipse through point D with the foci at points A and B and a circle with the center C and the radium $|CD|$. Write D' and D'' the intersections of CA and the ellipse and CB and the circle respectively. It is obvious that $|CD| < |CD'| < |CA|$. So

$$\begin{aligned} |AD| + |BD| + |CD| &\leq |AD'| + |BD'| + |D'C| \\ &= |AC| + |BD'| < |AC| + |AB| \\ &\leq 2|AB|. \end{aligned}$$

By the Lemma and the fact that the longest edge of the triangle Av_3v_2 in Fig. 5(b) is less than $\sqrt{2}$, we get $l(P_4^*) < 2\sqrt{2} < 3$.

Finally suppose that the four points of P_4^* are all on boundary of the unit square and there is at least one point on each edge. The convex hull of P_4^* is a quadrilateral or a degenerate quadrilateral which is a right triangle. Let $l_1 \leq l_2 \leq l_3 \leq l_4$ denote the edge lengths of the quadrilateral. Obviously $l_1 + l_2 + l_3 + l_4 \leq 4$. If $l_4 \geq 1$ then

$$4 \geq l_1 + l_2 + l_3 + l_4 \geq l_1 + l_2 + l_3 + 1,$$

i.e. $l_1 + l_2 + l_3 \leq 3$.

If $l_4 < 1$ then $l_1 + l_2 + l_3 < 3$. So $l_1 + l_2 + l_3 = 3$ if and only if $l_1 = l_2 = l_3 = l_4 = 1$. This implies that the location of P_4^* is Fig. 2(c).

§4 Conjectures for n = 5, 6, 7, 8, 9

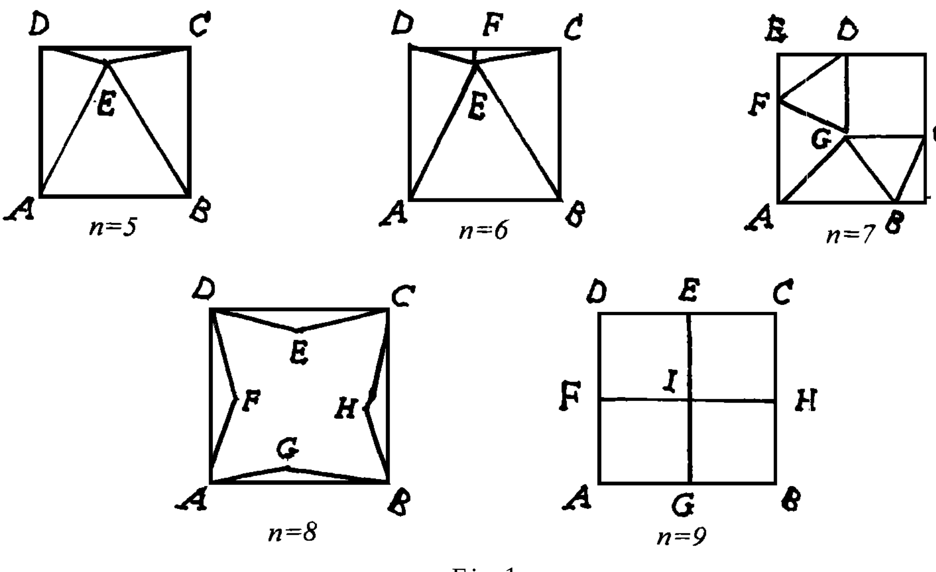

Fig. 1

We guess that the optimal locations of P_n^* ($n = 5, 6, 7, 8, 9$) as shown in Fig. 6.

$n=5, AE=BE=AB, C_5=2+3\sin(\pi/12)=3.0352276\cdots$;

$n=6, AE=BE=AB, EC=FC, C_6=3.13513\cdots$;

$n=7, DC=DE=GC=GE=GB=GF=FE=BC, C_7=3.33583\cdots$;

$n=8, AF=FG=GA=GB=GH=HB=HC=HE=EH=ED=DF=EF, C_8=3.62346631\cdots$.

$n=9$, E, F, G, H are the midpoints of CD, DA, AB, BC respectively and I is the crossing point with EG and HF, $C_9=4$.

The values of C_n ($n=2, 3, \cdots, 9$) are listed in Table 1.

Table 1

n	2	3	4	5	6	7	8	9
C_n	1.4142···	2.1180···	3	3.0352···	.1351···	3.3358···	3.6234···	4

§5. Conclusion and Remark

We have presented four general results for the whole construction on Maximum MST determined by n points in the unit square (see Theorem 1 to 4). But it is not enough to understand the location of P_n^*.

Some related problems remain open.

1) What is the asymptotic value of C_n?

2) Are the longest edges in optimal tree unique.

3) Does there exist an optimal tree in which the degree of the nodes is no more than 4?

4) How is a maximum MST determined by n points in other kinds of regions, such as unit circle, equilateral triangle ans so on?

References

[1] J. Schaer, The densest packing of 9 circles in a square, Canad Math. Bull 8(1965) 273- 277.

[2] G. Valette, A better packing of the equal circles in a square, North-Holland, Discrete Math. 76(1989) 57- 59.

[3] D. Z Du and F. K. Hwang, An new bound for the steiner ratio, Trans Amer. Math. Soc , 278(1983) 137- 148.

[4] Michel Gondran and Michek Minoux, Graphs and algorithms, John Wiley and Sons, 1984.