Information Processing Letters 70 (1999) 241–243

Information Processing Letters
www.elsevier.com/locate/ipl

On average edge length of minimum spanning trees

Suman Kumar Nath [1], Rezaul Alam Chowdhury [2], M. Kaykobad [3]

Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

Received 21 January 1999; received in revised form 25 March 1999
Communicated by F.B. Schneider

Abstract

This paper presents a theorem that asserts that average edge length of the minimum spanning tree of a complete graph on $n+1$ vertices is less than or equal to the average edge length of all the $n+1$ minimum spanning trees of the induced graph on n vertices. The result is also in compliance with results given by Frieze and Steele. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Minimum spanning tree; Complete graph; Average edge length

1. Introduction

Let K_n denote a complete graph with vertex set V with cardinality n and edge set E. Weights are assigned to each $e \in E$ by means of non-negative, independent random variables X_e with common distribution F. We further let $l(n)$ denote the cost of the minimum spanning tree of K_n, i.e.,

$$l(n) = \min_T \sum_{e \in T} X_e$$

where the minimum is taken over the set of all n^{n-2} trees which span V.

Under the assumption that the X_e are uniformly distributed on $[0, 1]$, Frieze [2] established that

$$\lim_{n\to\infty} E\bigl(l(n)\bigr) = \zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.202\ldots$$

and $l(n) \to \zeta(3)$ in probability 1 as $n \to \infty$.

In Frieze [2] the result was extended to cover the case of continuous F with finite variances. Steele [4] extended Frieze's theorem to the widest possible class of F. In particular, Steele gave the following result:

If X_e are independent non-negative random variables whose continuous distribution function F is differentiable from the right at 0, *with $F'(0) > 0$, then $l(n)$ converges to $\zeta(3)/F'(0)$ in probability* 1, *i.e., for all* $\varepsilon > 0$,

$$P\Bigl(\bigl|l(n) - \zeta(3)/F'(0)\bigr| > \varepsilon\Bigr) \to 0 \quad \textit{as } n \to \infty.$$

In this paper we present a related result which states how the average edge length of minimum spanning trees formed from a random complete graph varies with its number of vertices. In the next section, we give the result. We also show the relation of our result to Frieze and Steele in the conclusion.

2. The result

Let us consider a K_{n+1} with V as set of vertices, where $|V| = n+1$. Let $\Psi(n+1)$ be the average

[1] Email: sknath@bdonline.com.
[2] Email: shaikat@bdonline.com.
[3] Email: kaykobad@buet.edu.

PII: S0020-0190(99)00068-X

weight of edges in a minimum spanning tree of K_{n+1}. Removing any vertex from K_{n+1} gives a K_n. K_n can be induced from K_{n+1} in $(n+1)$ different ways. For each particular K_n, we can calculate $\Psi(n)$. Let $\overline{\Psi}(n)$ be the average of all the $\Psi(n)$ obtained in this way. Then we have the following result.

Theorem 2.1. $\Psi(n+1) \leqslant \overline{\Psi}(n)$.

Proof. Let $l(n+1)$ be the sum total of weights on n edges of the minimum spanning tree of K_{n+1}. Then, $\Psi(n+1) = l(n+1)/n$. There are $(n+1)$ ways of inducing a K_n from a K_{n+1}. Let $l_i(n+1)$ be the sum total of weights of the $(n-1)$ edges of the minimum spanning tree of K_n when K_n is constructed from K_{n+1} by removing vertex i. Then,

$$\overline{\Psi}(n) = \frac{1}{(n+1)} \sum_{i=1}^{n+1} \left(\frac{l_i(n+1)}{n-1} \right). \tag{2.1}$$

Let s_i be the smallest weight of the edges connected to vertex i. Then $l(n+1) \leqslant l_i(n+1) + s_i$. So,

$$\begin{aligned} \Psi(n+1) &= \frac{l(n+1)}{n} \\ &\leqslant \frac{1}{n(n+1)} \sum_{i=1}^{n+1} \big(l_i(n+1) + s_i \big) \\ &= \frac{1}{(n+1)} \left(\frac{\sum_{i=1}^{n+1} l_i(n+1)}{n-1} + \frac{\sum_{i=1}^{n+1} s_i}{n} \right. \\ &\qquad \left. - \frac{\sum_{i=1}^{n+1} l_i(n+1)}{n(n-1)} \right). \end{aligned} \tag{2.2}$$

Now, subtracting (2.1) from (2.2) yields

$$\begin{aligned} &\Psi(n+1) - \overline{\Psi}(n) \\ &\quad \leqslant \frac{1}{(n+1)} \left(\frac{\sum_{i=1}^{n+1} s_i}{n} - \frac{\sum_{i=1}^{n+1} l_i(n+1)}{n(n-1)} \right) \\ &\quad = \frac{1}{n(n+1)(n-1)} \left((n-1) \sum_{i=1}^{n+1} s_i \right. \\ &\qquad \left. - \sum_{i=1}^{n+1} l_i(n+1) \right). \end{aligned} \tag{2.3}$$

Now concentrate on the second term within the bracket. We have $(n+1)$ $T_i(n)$'s with n vertices, where $T_i(n)$ is a minimum spanning tree with of the graph with n vertices which is induced from K_{n+1} by removing vertex i $(i = 1, 2, 3, \ldots, n+1)$. Each of these $T_i(n)$'s has $(n-1)$ edges. We can systematically make a one-to-one mapping of each of these $(n-1)$ edges to $(n-1)$ vertices of $T_i(n)$. First we consider vertex $(i \bmod (n+1)) + 1$ as the root of $T_i(n)$. Then we start from the leaves and map the edge incident on a leaf to that leaf. Now removing the mapped leaves and edges from $T_i(n)$, we get a new tree and can apply the same mapping scheme, until all the $(n-1)$ edges map to unique vertices. Such a mapping scheme ensures that all the vertices of K_{n+1}, except vertex i (which is not included in $T_i(n)$), and $(i \bmod (n+1)) + 1$ (which is considered to be the root), have a one to one mapping to $(n-1)$ edges of $T_i(n)$. Suppose, in $T_i(n)$, the edge which maps to vertex j has weight d_{ij} with $j \neq i$, $j \neq (i \bmod (n+1)) + 1$. Clearly $d_{ij} \geqslant s_j$, since s_j is the minimum weight of the edges connected to vertex j. Let $N = \{1, 2, 3, \ldots, n+1\}$. So,

$$\begin{aligned} \sum_{i=1}^{n+1} l_i(n+1) &= \sum_{i=1}^{n+1} \sum_{\substack{j=1, j\neq i, \\ j \neq (i \bmod (n+1))+1}}^{n+1} d_{ij} \\ &\geqslant \sum_{i=1}^{n+1} \sum_{j \in N - \{(i \bmod (n+1))+1\}} s_j \\ &= (n-1) \sum_{j=1}^{n+1} s_j \\ \Rightarrow \left\{ (n-1) \sum_{i=1}^{n+1} s_i - \sum_{i=1}^{n+1} l_i(n+1) \right\} &\leqslant 0. \end{aligned} \tag{2.4}$$

From Eq. (2.3), it follows that $\Psi(n+1) \leqslant \overline{\Psi}(n)$. Hence the theorem is proved. □

This is also in compliance with the usual expectation that deleting a vertex from a graph will cause an increase in pairwise distance between the remaining vertices.

3. Conclusion

The result we have presented in the previous section does agree with the results of Frieze [2] and Steele [4]. Their results state that when the edge weights of a complete graph are identically and independently distributed from a distribution satisfying mild constraints, the total weight of the edges in a minimum spanning tree is asymptotically constant. This implies that if the number of nodes increases, average edge length decreases. Our theorem generalizes the result for any graph without any such constraints. The result is of theoretical importance and can be used as a bound in average edge length of spanning tree in further research.

References

[1] A.V. Aho, J.E. Hopcroft, J.D. Ullman, Data Structures and Algorithms, Addison-Wesley, Reading, MA, 1983.
[2] A.M. Frieze, On the value of a minimal spanning tree problem, Discrete Appl. Math. 10 (1985) 47–56.
[3] G.S. Lueker, Optimization problems on graph with independent random edge weights, SIAM J. Comput. 10 (1981) 338–351.
[4] J.M. Steele, On Frieze's $\zeta(3)$ limit for lengths of minimal spanning trees, Discrete Appl. Math. 18 (1987) 99–103.