IOAN TOMESCU

An upper bound for the shortest hamiltonian path in the symmetric euclidean case

AN UPPER BOUND FOR THE SHORTEST HAMILTONIAN PATH IN THE SYMMETRIC EUCLIDEAN CASE (*)

by Ioan Tomescu ([1])

Abstract. — *In this paper an algorithm for obtaining a Hamiltonian path from a shortest spanning tree of a complete weighted graph is proposed. As a consequence, two inequalities between the costs of a shortest Hamiltonian path and a shortest spanning tree T in the symmetric Euclidean case are proposed. These inequalities involve the diameter of T or the number of terminal vertices of T and they become equalities in some particular cases.*

Keywords: Hamiltonian path, Spanning tree, Triangle inequality.

Résumé. — *Dans ce travail on propose un algorithme pour l'obtention d'une chaîne hamiltonienne en partant de l'arbre minimal d'un graphe complet non orienté valué. Comme conséquence l'on déduit une borne du coût d'une chaîne hamiltonienne minimale dans le cas où les coûts satisfont l'inégalité du triangle. Cette borne fait intervenir le diamètre et le nombre des sommets pendants de T, elle devient une égalité dans certains cas particuliers.*

Mots clés : Chaîne hamiltonienne, arbre partiel, inégalité du triangle.

1. DEFINITIONS AND NOTATION

Let K_n be the complete graph having vertex set $V(K_n) = \{1, \ldots, n\}$ and edge set $E(K_n) = \{ij \mid 1 \leq i < j \leq n\}$.

To each edge ij is associated a cost $c(i,j) = c_{ij} \geq 0$ such that:

$$C = (c_{ij})_{i,j=1,\ldots,n}$$

is a symmetric matrix having $c_{ii} = 0$ on the main diagonal.

Suppose also that elements of C verify triangle inequality:

$$c_{ij} \leq c_{ik} + c_{kj}$$

for any $i, j, k = 1, \ldots, n$, i.e. c_{ij} may be considered as a distance between vertices i and j of K_n.

The cost of a path (cycle) of K_n is equal to the sum of the costs of the edges of this path (cycle).

A shortest spanning tree (SST) of K_n is a spanning tree of K_n having minimum cost. In a similar way a shortest Hamiltonian path (SHP) and a shortest Hamiltonian cycle (SHC) are defined.

The problem of determining a shortest Hamiltonian cycle of K_n is also known as the travelling salesman problem [4].

Some connections between SST's problem and SHC's problem were discovered by M. Held and R. M. Karp [5].

Using a similar method of penalizing vertices, N. Christofides proposed an algorithm for solving SHP's problem which starts from a SST of K_n [3]. This algorithm is not necessarily convergent and it produces a variety of spanning trees of K_n associated to matrices of costs preserving all SHP having two fixed extremities. I. Tomescu proposed an algorithm for obtaining a Hamiltonian path (HP) of K_n [7] which is based on 3 types of transformations of a SST, denoted respectively by 1, 2 and 3.

These transformations are used in order to decrease at each step the number of terminal vertices (or leaves) of the tree, until it is obtained a spanning tree of K_n having exactly two terminal vertices, which is a HP of K_n.

No estimation is made in [7] about the error of the algorithm for obtaining a SHP, even in the Euclidean symmetric case, when the costs satisfy triangle inequality.

In the sequel we shall obtain such an evaluation if only transformations of type 3 are applied.

For a spanning tree T of K_n we shall define the distance $d_T(i,j)$ between vertices i and j by the cost of the unique path of T joining i and j and the diameter of T, denoted by diam(T), as

$$\text{diam}(T) = \max_{1 \le i, j \le n} d_T(i,j)$$

We shall denote by c(SST), c(SHP), c(HP), c(SHC), the cost of a SST, SHP, HP, respectively SHC of K_n.

The degree of a vertex x in a tree T is denoted $\deg_T(x)$.

An elementary subdivision of a nonempty graph G is a graph obtained from G by the removal of some edge $e = uv$ and the addition of a new vertex w and edges uw and vw.

A subdivision of G is a graph obtained from G by a succession of elementary subdivisions [1].

2. UPPER BOUNDS FOR c (SHP)

Suppose now that T is a SST of K_n having $t(T) = t$ terminal vertices. If $t = 2$ then T is a SHP of K_n. Otherwise, $t \geq 3$ and we shall describe an algorithm which transforms T into a Hamiltonian path of K_n, denoted HP, such that:

$$c(\text{HP}) \leq 2 c(T) - \text{diam}(T).$$

For this, let i, j be two terminal vertices of T such that $d_T(i, j) = \text{diam}(T)$. If x is a terminal vertex of T, $x \neq i, j$, there exists a unique path $P : x, x_1, x_2, \ldots, x_p$ of T, $p \geq 1$, such that $\deg_T(x_1) = \deg_T(x_2) = \ldots = \deg_T(x_{p-1}) = 2$ and $\deg_T(x_p) \geq 3$. There is a vertex v adjacent to x_p such that $v \notin V(P)$.

From T we obtain two trees T_1 and T_1' defined as follows:

T_1 is obtained from T by deleting vertices $x, x_1, \ldots, x_{p-1}$ and their incident edges;

T_1' is obtained from T by deleting edge $x_p v$ and inserting new edge xv, hence $V(T_1') = V(T) = \{1, \ldots, n\}$.

It is clear that T_1' is isomorphic to a graph obtained from T_1 by inserting p new vertices on the edge $x_p v$, or T_1' is a subdivision of T_1 and $t(T_1) = t(T_1') = t - 1$.

We can write:

$$c(T_1') = c(T) - c(x_p, v) + c(x, v) \tag{1}$$

By applying triangle inequality we derive:

$$
\begin{aligned}
c(x, v) \leq c(x, x_1) + c(x_1, v) &\leq c(x, x_1) \\
&+ c(x_1, x_2) + c(x_2, v) \leq \ldots \leq c(x, x_1) + c(x_1, x_2) + \ldots \\
&+ c(x_{p-1}, x_p) + c(x_p, v) = d_T(x, x_p) + c(x_p, v).
\end{aligned}
$$

From (1) we deduce that:

$$c(T_1') \leq c(T) + d_T(x, x_p) = c(T) + c(P) \tag{2}$$

If $t = 3$ then $t(T_1) = t(T_1') = 2$ and we define HP $= T_1'$.

Otherwise, let y be a terminal vertex of T_1, $y \neq i, j$ and the path:

$$P_1 : \quad y, y_1, y_2, \ldots, y_q$$

of T_1, $q \geq 1$, such that:

$$\deg_{T_1}(y_1) = \deg_{T_1}(y_2) = \ldots = \deg_{T_1}(y_{q-1}) = 2 \text{ and } \deg_{T_1}(y_q) \geq 3.$$

Since T_1' is a subdivision of T_1 it follows that all terminal vertices and all vertices of degree greater than 2 are the same in T_1 and T_1'.

Hence we derive that there is a path Q_1 between y and y_q in T_1' such that all vertices of this path different from y and y_q have their degrees in T_1' equal to 2. But $\deg_{T_1'}(y_q) \geqq 3$ implies that there exists a vertex w adjacent to y_q in T_1' such that $w \notin V(Q_1)$. From T_1 we obtain a new tree T_2 and from T_1' a tree T_2' which are defined as follows:

T_2 is deduced from T_1 by the removal of vertices $y, y_1, \ldots, y_{q-1}$ and their incident edges and T_2' is obtained from T_1' by deleting edge $y_q w$ and inserting new edge yw, hence $V(T_2') = \{1, \ldots, n\}$.

T_2' is a subdivision of T_2 and $t(T_2) = t(T_2') = t - 2$.

We derive also:

$$c(T_2') = c(T_1') - c(y_q, w) + c(y, w). \tag{3}$$

By applying triangle inequality we get:

$$c(y, w) \leqq c(y, y_1) + c(y_1, w) \leqq \ldots$$
$$\leqq c(y, y_1) + c(y_1, y_2) + \ldots + c(y_{q-1}, y_q) + c(y_q, w)$$
$$= d_{T_1}(y, y_q) + c(y_q, w) = d_T(y, y_q) + c(y_q, w),$$

since T_1 is a subtree of T.

Hence (3) implies:

$$c(T_2') \leqq c(T_1') + d_T(y, y_q) = c(T_1') + c(P_1). \tag{4}$$

If $t = 4$ then T_2' is a Hamiltonian path and we define $HP = T_2'$. Otherwise, we consider a terminal vertex $z \neq i, j$ of T_2 and we obtain in a similar manner trees T_3 and T_3' and so on.

Note that T_i are subtrees of T containing all the $i-j$ path of T whereas T_i' are spanning trees of K_n, as is illustrated in figure 1.

It follows that T_{t-2} is the unique path P_{t-2} of T joining vertices i and j and T_{t-2}' is a Hamiltonian path of K_n. We define $HP = T_{t-2}'$. From inequalities (2), (4), ... we deduce that:

$$c(HP) = c(T_{t-2}') \leqq c(T) + \sum_{i=0}^{t-3} c(P_i),$$
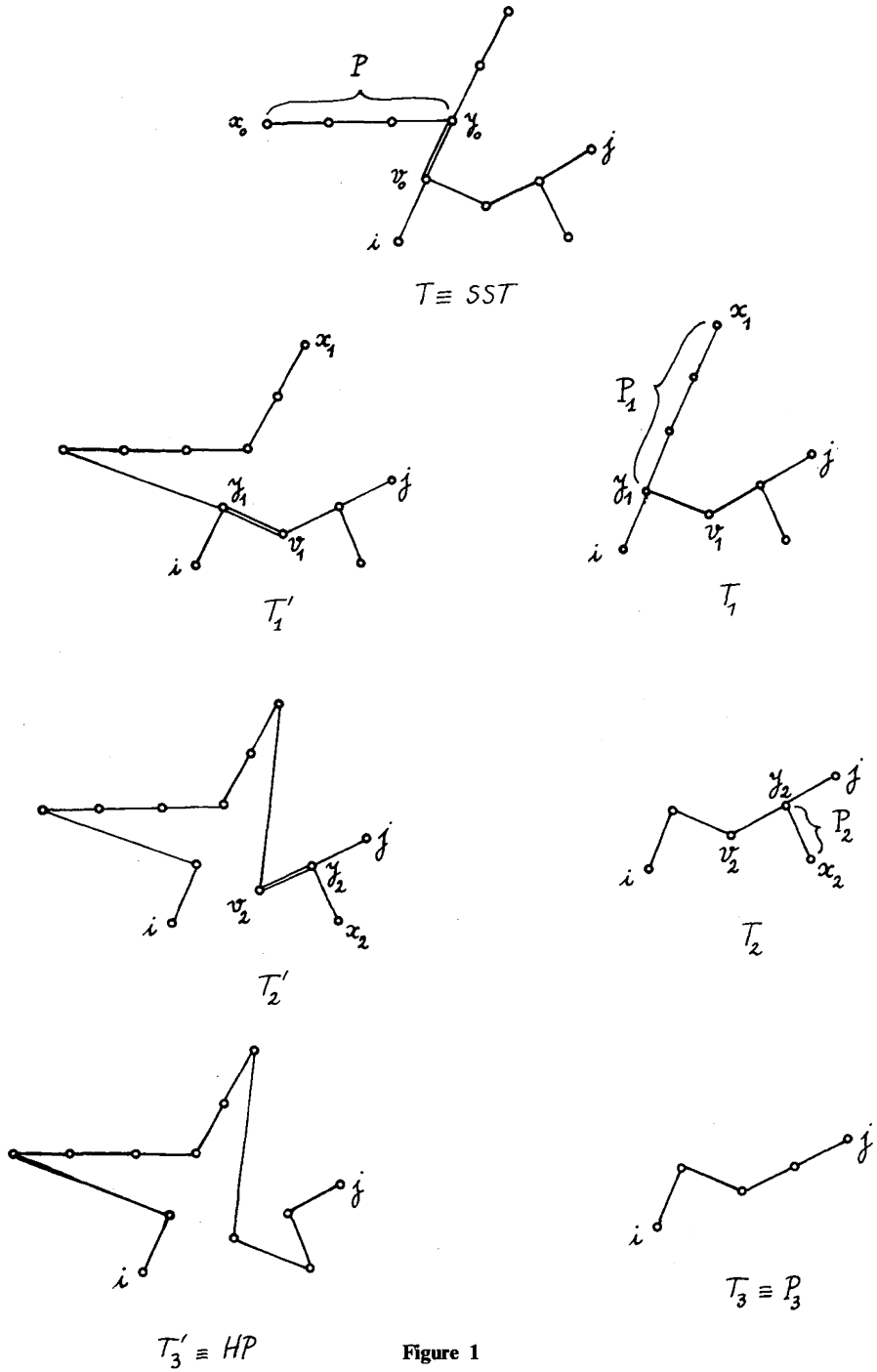
where $P_0 = P$.

Figure 1

Since $E(P) \cup E(P_1) \cup \ldots \cup E(P_{t-2})$ is a partition of the edge set $E(T)$ of T (see *fig.* 1), we can write:

$$\sum_{i=0}^{t-3} c(P_i) = \sum_{i=0}^{t-2} c(P_i) - c(P_{t-2}) = c(T) - d_T(i,j) = c(T) - \text{diam}(T),$$

hence:

$$c(\text{HP}) \leqq 2c(T) - \text{diam}(T). \tag{5}$$

This heuristic algorithm for obtaining a SHP may proceed as follows: at each step is selected that terminal vertex $x \neq i, j$ such that the increase of the cost, equal to $c(x,v) - c(x_p,v)$ to be minimum.

THEOREM 1: *If* HP *is a Hamiltonian path obtained with the preceding algorithm in the symmetric Euclidean case from T, which is a SST having t terminal vertices, then:*

$$c(\text{HP}) \leqq 2c(T) - \text{diam}(T) \leqq 2\left(1 - \frac{1}{t}\right)c(T). \tag{6}$$

Proof: It remains to prove only that:

$$2c(T) - \text{diam}(T) \leqq 2\left(1 - \frac{1}{t}\right)c(T), \tag{7}$$

or $\text{diam}(T) \geqq (2/t)\,c(T)$.

For this consider a terminal vertex $a \neq i, j$ of T, where $d_T(i,j) = \text{diam}(T)$. There exists a unique path $a, a_1, a_2, \ldots, a_r, b$ in T such that b is the unique vertex of this path which belongs to the path between i and j in T (see *fig.* 2).
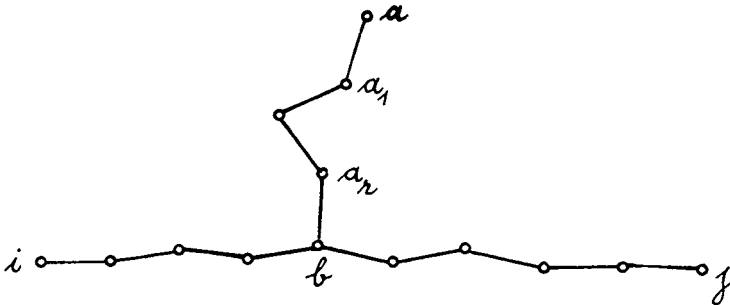

Figure 2

Since $d_T(i,j) = \text{diam}(T)$ it follows that:

$$d_T(a,b) \leqq d_T(i,b) \qquad \text{and} \qquad d_T(a,b) \leqq d_T(j,b),$$

hence:

$$2\,d_T(a,b) \leqq d_T(i,b) + d_T(b,j) = \text{diam}(T),$$

or:

$$d_T(a,b) \leqq \frac{1}{2}\,\text{diam}(T).$$

Any edge not belonging to the path between i and j in T joins two consecutive vertices of a path of the form $a, a_1, \ldots, a_r, b$ in T, where $a \neq i, j$ is a terminal vertex of T and b lies on the $i-j$ path in T. This implies:

$$c(T) \leqq d_T(i,j) + \sum_{a \neq i,j} d_T(a,b)$$

$$\leqq \text{diam}(T) + \frac{t-2}{2}\,\text{diam}(T) = \frac{t}{2}\,\text{diam}(T),$$

or:

$$\text{diam}(T) \geqq \frac{2}{t}\,c(T)$$

and (7) is proved.

COROLLARY 1: *If SHP is a shortest Hamiltonian path in the symmetric Euclidean case and T is a SST with t terminal vertices, the following inequalities hold:*

$$c(\text{SHP}) \leqq 2\,c(T) - \text{diam}(T) \leqq 2\left(1 - \frac{1}{t}\right)c(T). \tag{8}$$

If the star $K_{1,n-1}$ is a spanning tree of K_n such that the cost of any edge of $E(K_{1,n-1})$ is equal to 1 and the cost of any edge of $E(K_n) \backslash E(K_{1,n-1})$ is equal to 2, the triangle inequality is verified.

In this case $K_{1,n-1}$ is a SST of K_n, $c(\text{SHP}) = 2n-4$, $c(K_{1,n-1}) = n-1$, $\text{diam}(K_{1,n-1}) = 2$, $t = n-1$ and all inequalities of (8) become equalities. This case is shown in figure 3.

Now consider the spanning tree T of K_n composed from the path $1, 2, \ldots, n-t+2$ and $t-2$ edges: $2, n-t+3$; $3, n-t+4$; $\ldots$; $t-1, n$, where $n \geqq 2(t-1)$, $t \geqq 4$ such that the cost of any edge of $E(T)$ is equal to 1 and the

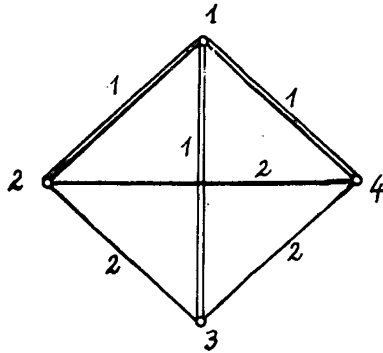cost of any edge $uv \in E(K_n) \setminus E(T)$ is equal to $d_T(u, v)$. An illustration is given in figure 4.
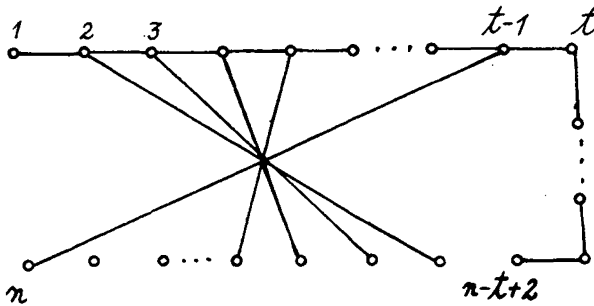

Figure 3

Figure 4

In this case the triangle inequality is satisfied, T is a SST of K_n having $t(T) = t$ terminal vertices and $c(T) = n - 1$, $\text{diam}(T) = n - t + 1$. For a SHP of K_n at least $t - 2$ terminal vertices of T are internal vertices, hence:

$$c(\text{SHP}) \geqq c(T) + t - 2 = n + t - 3.$$

It follows that we can choose:

SHP: 1, 2, $n-t+3$, 3, $n-t+4, \ldots, t-1$, n, t, $t+1, \ldots, n-t+2$, and $c(\text{SHP}) = n + t - 3$.

We obtain:

$$c(\text{SHP}) = 2c(T) - \text{diam}(T) < 2\left(1 - \frac{1}{t}\right)c(T)$$

since the inequality is equivalent to $n > t + 1$, which is true because $n \geq 2t - 2$ and $t \geq 4$.

Note that if the proposed algorithm is applied to T we find a SHP of K_n.

COROLLARY 2: *If HP is a Hamiltonian path produced by the proposed algorithm in the symmetric Euclidean case from a SST having t terminal vertices, it follows that:*

$$\frac{c(\text{HP})}{c(\text{SHP})} \leq 2\left(1 - \frac{1}{t}\right).$$

This inequality results from (6) since $c(\text{SST}) \leq c(\text{SHP})$.

D. Rosenkrantz, R. Stearns and P. Lewis [6] proved that in the symmetric Euclidean case if T is a SST of K_n then $c(\text{SHC}) \leq 2c(T)$ (*see* also [2]). This result appears also as a consequence of the theorem 1:

COROLLARY 3 (Rosenkrantz, Stearns, Lewis): *If the costs satisfy triangle inequality then:*

$$c(\text{SHC}) \leq 2c(T),$$

where T is a SST of K_n.

Proof: If HP is a Hamiltonian $i-j$ path obtained with the proposed algorithm, then adding edge ij we find a HC for which the cost is equal to:

$$c(i,j) + c(\text{HP}) \leq d_T(i,j) + 2c(T) - \text{diam}(T) = 2c(T).$$

It follows that $c(\text{SHC}) \leq 2c(T)$.

COROLLARY 4: *If HC denotes the Hamiltonian cycle obtained in the symmetric Euclidean case from the Hamiltonian $i-j$ path HP by adding edge ij, then:*

$$c(\text{HC}) \leq 2c(T) \leq 2(c(\text{SHC}) - c(e_{\max})),$$

where T is the SST from which HP is deduced and $e_{\max}$ denotes an edge of SHC having maximum cost.

Proof: By the removal of the edge e_{max} of SHC we get a HP, which is a spanning tree of K_n, hence:

$$c(T) \leq c(SHC) - c(e_{max}).$$

Note that this upper bound for the cost of HC is the same as for the Hamiltonian cycle obtained by the nearest insertion rule (NIR procedure of [4]), proposed by D. Rosenkrantz, R. Stearns and P. Lewis [6].

From corollary 4 it follows also that:

$$\frac{c(HC)}{c(SHC)} \leq 2 - \frac{1}{n}$$

since:

$$c(e_{max}) \geq \frac{c(SHC)}{n}.$$

ACKNOWLEDGEMENTS

REFERENCES

1. M. Behzad, G. Chartrand and L. Lesniak-Foster, *Graphs and Digraphs*, Prindle, Weber and Schmidt, Boston, Massachusetts, 1979.
2. J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, The Macmillan Press, London, 1977.
3. N. Christofides, *The Shortest Hamiltonian Chain of a Graph*, S.I.A.M. J. Appl. Math., Vol. 19, 1970, pp. 689-696.
4. N. Christofides, *The Travelling Salesman Problem*, Combinatorial Optimization, John Wiley, 1979, pp. 131-149.
5. M. Held and R. M. Karp, *The Travelling-Salesman Problem and Minimum Spanning Trees*, Operations Research, Vol. 18, 1970, pp. 1138-1162.
6. D. J. Rosenkrantz, R. E. Stearns and P. M. Lewis, *An Analysis of Several Heuristics for the Traveling Salesman Problem*, S.I.A.M. J. Comput., Vol. 6, No. 3, 1977, pp. 563-581.
7. I. Tomescu, *Un algorithme pour l'obtention d'une chaîne hamiltonienne en partant de l'arbre minimal d'un graphe*, R.A.I.R.O., V-3, Vol. 9, 1975, pp. 5-12.