The TSP and the Sum of its Marginal Values*

Moshe Dror† Yusin Lee‡ James B. Orlin§ Valentin Polishchuk¶

December 20, 2005

Abstract

This paper introduces a new notion related to the traveling salesperson problem (TSP) — the notion of the *TSP ratio*. The TSP ratio of a TSP instance I is the sum of the marginal values of the nodes of I divided by the length of the optimal TSP tour on I, where the marginal value of a node $i \in I$ is the difference between the length of the optimal tour on I and the length of the optimal tour on $I \setminus i$. We consider the problem of establishing exact upper and lower bounds on the TSP ratio. To our knowledge, this problem has not been studied previously.

We present a number of cases for which the ratio is never greater than 1. We establish a tight upper bound of 2 on the TSP ratio of any *metric* TSP. For the TSP on six nodes, we determine the maximum ratio of 1.5 in general, 1.2 for the case of metric TSP, and 10/9 for the geometric TSP in the L_1 metric. We also compute the TSP ratio experimentally for a large number of random TSP instances on small number of points.

Keywords: The traveling salesperson problem; mixed integer programming; cost allocation.

1 Introduction

An instance of the traveling salesperson problem (TSP) on n nodes is specified by a complete undirected graph $G = (V, E)$ with $V = \{1 \dots n\}$ and a *cost*, c_{ij}, for every edge $(i, j) \in E$. The TSP asks one to compute a minimum-cost circuit in G that visits each node exactly once. The problem has been extensively studied in combinatorial optimization; see, e.g., [2].

We assume that $c_{ij} \geq 0$ and that $c_{ij} = c_{ji}$ for all $i, j \in V$. In a *metric* TSP the costs, in addition, obey the *triangle inequality*: $c_{ij} + c_{jk} \geq c_{ik}$ for all $i, j, k \in V$. In a *geometric* TSP a node $i \in V$ is embedded in the plane, at point (x_i, y_i), and the cost c_{ij} is given by the L_2

*Preliminary version [1] appeared in the abstracts of the *14th Annual Fall Workshop on Computational Geometry*, November 2004.

†mdror@eller.arizona.edu. MIS Department, College of Business and Public Administration, University of Arizona Tucson, Arizona 85721, USA. Work done while M. Dror was visiting Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

‡yusin@mail.ncku.edu.tw. Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue Cambridge, Massachusetts, 02139, USA. Now at Department of Civil Engineering, National Cheng Kung University Tainan, Taiwan

§jorlin@mit.edu. Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue Cambridge, Massachusetts, 02139, USA. This work was supported in part by Office of Naval Research grant ONR N00014-98-1-0317.

¶valentin.polishchuk@stonybrook.edu. Applied Mathematics and Statistics Department, Stony Brook University Stony Brook, New York, 11794-3600, USA. V. Polishchuk is partially supported by the National Science Foundation (CCF-0431030), NASA Ames (NAG2-1620), and a grant from Metron Aviation.

distance $c_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. We consider also distances measured by the L_p metric: $c_{ij} = [|x_i - x_j|^p + |y_i - y_j|^p]^{1/p}$.

For a given instance I of the TSP, we let $TSP(I)$ denote the cost of an optimal solution of I. We define the *marginal cost of point* i, m_i, to be $m_i = TSP(I) - TSP(I \setminus i)$. Let the *TSP ratio* of I be $f(I) = \sum_i m_i / TSP(I)$, the sum of the marginal costs normalized by $TSP(I)$.

In this paper we study the relationship between $TSP(I)$ and $\sum_{i=1}^n m_i$.

We are motivated to study this question for at least two reasons. First, it examines a fundamental relationship for a well-studied combinatorial optimization problem. Second, the relationship between the marginal costs and the optimal cost $TSP(I)$ is interesting from a game theoretic point of view. It may be reasonable for a customer at node i to be expected to pay a cost related to the marginal cost m_i. If the sum of the marginals, $\sum_{i=1}^n m_i$, is larger than $TSP(I)$, then the cost of the tour can be covered by allocating cost to the customers according to their marginal costs. If $m_i < 0$, so that it is advantageous for a player at node i to be added to the tour, then the other players are motivated to entice player i to join in order to reduce the cost. Our work was motivated by Dror and Ferland ([3]), who have studied the TSP from a game theoretic perspective, in which players are asked to pay for the tour according to their marginal costs m_i.

1.1 Results

- In Sec. 2 we provide some general upper and lower bounds on the TSP ratio. We prove that any instance I on $n \leq 5$ points has $f(I) \leq 1$, and provide examples of instances I on $n = 6$ points for which $f(I) > 1$. We show that in the case of the *metric* TSP, the TSP ratio is bounded from above by 2, and that this bound is tight. We also outline an important property of an instance of the *geometric* TSP in the L_1 metric with maximum TSP ratio.

- In Sec. 3 we consider the case $n = 6$. We formulate the problem of finding an instance I with the maximum value of $f(I)$ as a mixed integer program. By solving the program we obtain the instance and the value $f(I) = 1.5$. Next, we impose the triangle inequality as a set of additional constraints in the program and obtain the instance with the maximum value of TSP ratio, 1.2, over all *metric* TSP instances. Finally, we propose the program for finding the instance I of *geometric* TSP in the L_1 metric with the highest possible value of $f(I) = 10/9$.

- In Sec. 4 we present the results of finding the TSP ratio for $36 \cdot 10^9$ geometric TSP instances in different metrics on $n = 6 \ldots 11$ points. We report that $f(I) > 1$ occurred very rarely in our experiments.

2 Preliminaries

In this section we make some general observations about the TSP ratio.

Definition 2.1. *A TSP instance I is said to have the fixed order property if there exists an optimal TSP tour $\mathbb{S}$ on I, such that for all $i \in I$, $\mathbb{S} \setminus i$ is optimal for $I \setminus i$.*

Clearly, any geometric TSP instance with the points in convex position has this property.

Theorem 2.2. *If an instance I of the TSP satisfies the fixed order property, then $f(I) \leq 1$.*

Proof. Assume without loss of generality that 1-2-…-n is the optimal TSP tour. In this case the marginal value of a node $i \in I$, $m_i = c_{i-1,i} + c_{i,i+1} - c_{i-1,i+1}$ (we assume arithmetic modulo n in the indices throughout). Then

$$\begin{aligned} \sum_i m_i - TSP(I) &= TSP(I) - \sum_i c_{i-1,i+1} = \\ &= c_{12} + c_{23} + \cdots + c_{n1} - c_{n2} - c_{13} - \cdots - c_{n-1,1} \end{aligned}$$

Since 1-2-…-n is the optimal TSP tour, by the local optimality $c_{12} + c_{34} \leq c_{13} + c_{24}$, $c_{23} + c_{45} \leq c_{24} + c_{35}$, …, $c_{n-1,n} + c_{12} \leq c_{n-1,1} + c_{n2}$, $c_{n1} + c_{23} \leq c_{n2} + c_{13}$. Summing up and dividing by 2 implies $c_{12} + c_{23} + \cdots + c_{n1} \leq c_{n2} + c_{13} + \cdots + c_{n-1,1}$, which implies $\sum_{i=1}^{n} m_i \leq TSP(I)$, or $f(I) \leq 1$. □

Corollary 2.3. *Any geometric TSP instance I with the points in convex position has $f(I) \leq 1$.*

Corollary 2.4. *Any instance I of TSP on $n \leq 4$ points has $f(I) \leq 1$.*

Proof. The corollary is trivially true for $n \leq 3$. Let $\mathbb{S}$ =1-2-3-4 be the optimal tour for an instance I on 4 nodes. For any $i \in \{1 \dots 4\}$, $\mathbb{S} \setminus i$ is the optimal tour on $I \setminus i$; thus I has the fixed order property. Hence, $f(I) \leq 1$. □

The question of whether or not the inequality $f(I) \leq 1$ holds in general was circulated for some time before the following theorem, showing that it does for all instances of size $n = 5$, and counterexamples of size $n = 6$ were obtained by Karl Juhnke, Joe Mitchell, and colleagues at Stony Brook ([4]).

Theorem 2.5. *Any instance I of TSP on 5 points has $f(I) \leq 1$.*

Proof. Assume without loss of generality that $\mathbb{S}$ =1-2-3-4-5 is the optimal TSP tour on I. For $i \in I$ call the tour $\mathbb{S} \setminus i$ on $I \setminus i$ the *canonical* tour.

Consider the instance $I \setminus 5$. Assuming that the tour starts at the node 1, there are only three possible tours on the nodes in the instance (up to reversing the tours directions): 1-2-3-4, 1-2-4-3 and 1-3-2-4. The first one is the canonical tour $\mathbb{S} \setminus 5$, the last one can not be (the only) optimal tour given that $\mathbb{S}$ is optimal on I. Thus, the only two possible optimal TSP sequences on $I \setminus 5$ are the canonical one and the second, non-canonical one. Similarly, there are only two possible optimal TSP tours on $I \setminus i$ for any $i = 1 \dots 4$. If all of these tours are canonical, then $f(I) \leq 1$ holds by Theorem 2.2. Thus, without loss of generality we may suppose that the non-canonical tour is optimal, say, on $I \setminus 5$. This leaves $2^4 = 16$ possibilities for the optimal tours on $I \setminus i, i = 1 \dots 4$. In each of the possibilities, it is easily shown that $f(I) \leq 1$.

For example, if the optimal tours on $I \setminus i, i = 1 \dots 5$ are 2-4-5-3, 1-3-4-5, 1-2-5-4, 1-3-2-5 and 1-2-4-3, then $\sum m_i - TSP(I) = 2TSP(I) - c_{34} - c_{52} - c_{13} - c_{45} - c_{25} - c_{14} - c_{14} - c_{13} - c_{25} - c_{24} - c_{31} = (c_{12} + c_{34} - c_{13} - c_{24}) + 2(c_{23} + c_{51} - c_{25} - c_{13}) + (c_{12} + c_{45} - c_{25} - c_{14}) \leq 0$ by the optimality of $\mathbb{S}$. In each of the other possibilities $\sum m_i \leq TSP(I)$ may be shown by the similar calculations. □

Example 2.6. *Consider the 6 points (0, 80), (30, 40), (0, 0), (42, 40), (72, 0), (72, 80) in the Euclidean plane. For this instance, $TSP(I) = 344$, while $\sum_i m_i = 360$; so the TSP ratio $f(I) = 1.0465$. In fact, the ratio can be increased somewhat by perturbing this instance, and adding two more points: (0, 0), (0.868, 0), (0, 1), (0.868, 1), (0.434, 0.595), (0.434, 0.405), (0.091, 0.5), (0.778, 0.5). For this instance with $n = 8$, the ratio is $f(I) \approx 1.054$.*

Example 2.7. *Let T be the equilateral triangle with base 2 and height $\sqrt{3}$. If I is the TSP instance on the vertices of the right triangular prism of height 1.5 with the base T, then $f(I) \approx 1.1$. This shows that Corollary 2.3 is not true in $\mathbb{R}^3$.*

In general, the TSP ratio is unbounded, as the following example demonstrates.

Example 2.8. *Recall (see e.g., [5, p. 61]), that a graph $G = (V, E)$ is called hypohamiltonian if G is not hamiltonian, but $G \setminus i$ is hamiltonian for any $i \in V$. It is known ([6]) that there exist hypohamiltonian graphs on n nodes for every $n \geq 13$ with possible exceptions of $n = 14, 17, 19$. Given a hypohamiltonian graph $G = (V, E)$ on n nodes, let I be the instance of TSP on V with $c_{ij} = 1$ for $(i, j) \in E$ and $c_{ij} = M$ otherwise, where M is some large number. Then $TSP(I) = n - 1 + M$, while $TSP(I \setminus i) = n - 1$ for all $i \in V$. Thus, $f(I) = \frac{nM}{n-1+M}$, which is arbitrarily close to n for large enough M.*

We remark that the instance from the above example is *not* an instance of metric TSP for $M > 2$. In the remainder of the paper, unless stated otherwise, we consider only metric TSP instances.

2.1 Metric TSP

Contrary to Example 2.8,

Theorem 2.9. *For any instance I of metric TSP, $f(I) \leq 2$.*

Proof. Assume without loss of generality that 1-2-...-n is the optimal TSP tour on I. For $i = 1 \ldots n$, consider a tour on $I \setminus i$; let j be the node, following $i-1$ on the tour. Adding the edges $(i-1, i)$ and (i, j) to the tour, converts it to a tour on I at the expense of $c_{i-1,i} + c_{ij} - c_{i-1,j} \leq 2c_{i-1,i}$ (note, that the triangle inequality is used here). Thus, $TSP(I) \leq TSP(I \setminus i) + 2c_{i-1,i}$, or $m_i \leq 2c_{i-1,i}$. Summing up this for $i = 1 \ldots n$, $\sum_i m_i \leq 2 \sum_i c_{i-1,i} = 2TSP(I)$. □

On the other hand, the TSP ratio can be arbitrarily close to 2. To see this just take $M = 2$ in the instance in Example 2.8.

The question of what the maximum value of the TSP ratio could be for an instance of the *geometric* TSP (under various metrics) is one of the most intriguing open problems in our study. In the next section we settle this questions for instances on $n = 6$ nodes in the L_1 (and the L_∞) metric.

Observe that in high-dimensional spaces, under the L_∞ metric, the TSP ratio can get arbitrarily close to 2. Indeed, by the Frechét Theorem, any (finite) metric space (X, d_X) with n elements can be embedded into $\mathbb{R}^n$ under the L_∞ metric; the embedding is $x \mapsto [d_X(x, i)]_{i \in X}$. In particular, the hypohamiltonian graph from Example 2.8 with $M = 2$ (which is a metric space) can be embedded into $\mathbb{R}^n$ under the L_∞ metric. For n large enough the TSP ratio of the instance, $\frac{nM}{n-1+M} = \frac{2n}{n+1}$, gets arbitrarily close to 2.

2.2 Geometric TSP in the L_1 Metric

Since the primary focus of the next section (and this paper in whole) is on geometric TSP instances in the L_1 metric, in the remainder of the paper, unless stated otherwise, we consider only such TSP instances. In this subsection we outline some properties of the TSP instance with maximum

TSP ratio. Although the results for the TSP on $n = 6$ points, presented in the next section, were obtained *without* using the observations that we make here, we remark (in the end of the next section) that taking Proposition 2.11 into account substantially decreases the number of the cases which need to be considered for $n = 6$, and provides a possibility to resolve the case $n = 7$.

For a set $S \subset \mathbb{R}^2$ let $BB(S)$ be the minimum axis-aligned (closed) bounding box of S. Let I be an instance of TSP with $f(I) > 1$. Suppose that there is a point $p \in I$ such that $p \notin BB(I \setminus p)$. We show that there exists a TSP instance I' with $f(I') > f(I) > 1$. The instance I' is obtained from I by replacing p with its projection p' onto $BB(I \setminus p)$: $I' = I \setminus p \cup p'$.

Lemma 2.10. $f(I') > f(I)$.

Proof. When passing through p', the TSP tour on I' may be extended to the TSP tour on I at the cost of $2d(p)$, where $d(p)$ is the distance from p to $BB(I \setminus p)$. Thus, $TSP(I) \leq TSP(I') + 2d(p)$.

On the other hand, consider a TSP tour on I. Since the tour passes through p, it must cross $BB(I \setminus p)$ twice — to reach p and to get back. We may modify the tour without increasing its length so that it passes through p'. After "clipping off" the part of the tour outside $BB(I \setminus p)$, we are left with a tour through I'. Since the length of the clipped part is $2d(p)$, we have $TSP(I') \leq TSP(I) - 2d(p)$. Thus, $TSP(I') = TSP(I) - 2d(p)$.

Similarly, for $i \in I \setminus p$, $TSP(I' \setminus i) = TSP(I \setminus i) - 2d(p)$. Obviously, $TSP(I \setminus p) = TSP(I' \setminus p')$. Hence,

$$f(I') = \frac{\sum_1^n (TSP(I') - TSP(I' \setminus i))}{TSP(I')} = \frac{\sum_1^n (TSP(I) - TSP(I \setminus i)) - 2d(p)}{TSP(I) - 2d(p)} =$$

$$= \frac{f(I)\,TSP(I) - 2d(p)}{TSP(I) - 2d(p)} > \frac{f(I)\,TSP(I) - f(I)\,2d(p)}{TSP(I) - 2d(p)} = f(I)$$

where the strict inequality holds provided $d(p) > 0$ since $f(I) > 1$. □

Thus,

Proposition 2.11. *If I is an instance of the geometric TSP in the L_1 metric with maximum TSP ratio, each side of $BB(I)$ has at least 2 points of I, where a corner point of $BB(I)$ is considered as an element of two sides.*

3 Maximum Ratio for Instances on $n = 6$ Nodes

In this section, we present a mixed integer linear program that formulates the problem of finding a TSP instance I on $n = 6$ nodes with the maximum TSP ratio as an optimization problem. We first consider the general case, without restricting ourselves to metric or geometric TSP instances; the decision variables in the program are thus the costs c_{ij}, $i, j \in I$, $i \neq j$.

Without loss of generality we assume that each tour starts in the node 1 (a tour on $I \setminus 1$ starts in 2). Then the number of different (up to reversing the direction) tours on I is $N = \frac{(n-1)!}{2} = 60$, and the number of different tours on $I \setminus i$, $i \in I$, is $N_1 = \frac{(n-2)!}{2} = 12$.

The other notation we use is as follows.

- For $k = 1 \ldots N$, let T_k^0 be the k^{th} tour on I. Let $T_1^0 =$ 1-2-3-4-5-6. We assume without loss of generality that this is the optimal tour, and that its length $TSP(I) = 1$.

i	1	2	3	4	5	6
1	0	0	1	0.2	0.5	0.2
2	0	0	0.3	0.5	0.8	0
3	1	0.3	0	0	0.8	0
4	0.2	0.5	0	0	0.5	0.2
5	0.5	0.8	0.8	0.5	0	0
6	0.2	0	0	0.2	0	0

i	1	2	3	4	5	6
1	0	0.1	0.3	0.2	0.3	0.2
2	0.1	0	0.2	0.3	0.2	0.3
3	0.3	0.2	0	0.1	0.2	0.3
4	0.2	0.3	0.1	0	0.3	0.2
5	0.3	0.2	0.2	0.3	0	0.1
6	0.2	0.3	0.3	0.2	0.1	0

Table 1: The cost matrices of the instances with the maximum ratios. Left: general case, right: the metric TSP case.

- For $i \in I$, $k = 1 \dots N_1$, let T_k^i be the k^{th} tour on $I \setminus i$. Let z_k^i be 1 if T_k^i is the minimum length tour on $I \setminus i$ and 0 otherwise.
- For $i \in I$, let $l^i = TSP(I \setminus i)$.

Our maximization problem can then be stated as follows:

$$
\begin{array}{lll}
\max & 6 - \sum_{i \in I} l^i & \\
\text{s.t.} & \sum_{(i,j) \in T_1^0} c_{ij} = 1 & \\
 & \sum_{(i,j) \in T_k^0} c_{ij} \geq 1 & k = 2 \dots N \\
l^i & \geq \sum_{(i,j) \in T_k^i} c_{ij} - 10(1 - z_k^i) & i \in I, k = 1 \dots N_1 \\
 & \sum_{k=1}^{N_1} z_k^i = 1 & i \in I \\
 & c_{ij} = c_{ji} & i, j \in I, i \neq j \\
 & 0 \leq c_{ij} \leq 1 & i, j \in I, i \neq j \\
 & z_k^i \in \{0, 1\} & i \in I, k = 1 \dots N_1
\end{array}
$$

It is a mixed integer programming problem, which is hard to solve directly as is. However, the following observation allows one to reduce the number of possible positive variables z_k^i from $6 \cdot 12$ to $6 \cdot 7$.

Observation. Given that T_1^0 is the optimal TSP tour on I, due to the local optimality, none of the five tours 2-5-4-3-6-2, 2-4-3-5-6-2, 2-3-5-4-6-2, 2-4-5-3-6-2 and 2-5-3-4-6-2 could be the only optimal tour on $I \setminus 1$. Similarly, for each $i \in \{2 \dots 6\}$, the are five tours that cannot be the only optimal tours on $I \setminus i$.

Thus, we may set the corresponding variables z_k^i to 0. After this is done (and the corresponding constraints are eliminated), the program can be solved with off the shelf software (for instance CPLEX) within a few minutes. The maximum TSP ratio of 1.5 is attained, e.g., on the cost matrix in Table 1, left.

3.1 Metric TSP

Observe that the TSP instance with the ratio of 1.5, found in the previous subsection, is not an instance of metric TSP. We can limit our search for an instance I with maximum $f(I)$ value to the instances of metric TSP by adding to our program a set of constraints $c_{ij} + c_{jk} \geq c_{ik}$, $i, j, k \in I$, imposing the triangle inequality on the costs. The program with the constraints added has a maximum of 1.2, with the costs given in the Table 1, right.

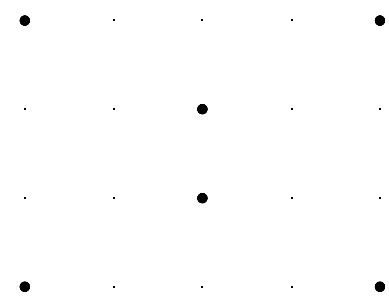

Figure 1: The 6-node geometric TSP in L_1 metric with the maximum TSP ratio of 10/9.

3.2 Geometric TSP in the L_1 Metric

Observe that the metric TSP instance with the ratio of 1.2, found in the previous subsection, is not an instance of geometric TSP (otherwise, the points 1,2,3 and 2,3,4 had to lie on the same line and $c_{14} = .2$ could not hold). We can modify our program in order to limit the search for an instance I with maximum $f(I)$ value to the instances of geometric TSP in the L_1 metric. The decision variables of the program become the coordinates (x_i, y_i) of the points $i = 1 \ldots 6$, and more additional variables and constraints are introduced as follows.

For $i, j \in I$, $i \neq j$, let $c^x_{ij} = |x_i - x_j|$, $c^y_{ij} = |y_i - y_j|$; let δ^x_{ij} be 1 if $x_i < x_j$ and 0 otherwise; let δ^y_{ij} be 1 if $y_i < y_j$ and 0 otherwise. The requirement that c_{ij} is given by the L_1 distance, $|x_i - x_j| + |y_i - y_j|$, can be modeled by the following set of constraints for all $i, j \in I, i \neq j$:

$$c_{ij} = c^x_{ij} + c^y_{ij} \qquad\qquad 0 \le x_i, y_i \le 1 \qquad\qquad \delta^x_{ij}, \delta^y_{ij} \in \{0, 1\}$$

$$\begin{array}{ll} x_i - x_j \le c^x_{ij} \le x_i - x_j + 2\delta^x_{ij} & \qquad x_j - x_i \le c^x_{ij} \le x_j - x_i + 2(1 - \delta^x_{ij}) \\ y_i - y_j \le c^y_{ij} \le y_i - y_j + 2\delta^y_{ij} & \qquad y_j - y_i \le c^y_{ij} \le y_j - y_i + 2(1 - \delta^y_{ij}) \quad . \end{array}$$

When the constraints are added, the program solution of 10/9 is found with the points placed in the plane as shown in Fig. 1.

We attempted to repeat the above experiment to establish the maximum TSP ratio for the instances on $n = 7$ points, but the solver did not obtain an optimal solution within several days of computation.

Remark. *By Corollary 2.3 and Proposition 2.11, if I is a TSP instance on $n = 6$ points with the maximum $f(I)$ value, there could only be either exactly one or exactly two points of I, lying strictly inside $BB(I)$.*

If there is only one such point, say point 6, then the optimal TSP tour on I may go only (up to a renumbering of the points) as shown in one of the plots in Fig. 2, left. In each of the cases, $\delta^{\cdot}_{ij}$ are known for $i, j = 1 \ldots 5$, and $\delta^{\cdot}_{6j}$ are known for all but two j from $1 \ldots 5$ (in fact, $\delta^x_{6\cdot}$ is known for one of them, and $\delta^y_{6\cdot}$ — for the other); this leaves only three (up to a reflection) possibilities for $\delta^{\cdot}_{6\cdot}$. Altogether, there remains only nine different cases to consider.

If there are two points, say p and k, strictly inside the bounding box of the others, the optimal TSP tour on I may go only (up to a renumbering of the points) as shown in one of the plots in Fig. 2, right. In each case, taking into account the optimality of the tour, there is at most two (up to the rotation or reflection of the whole scene) possibilities for δ^x_{kl}. For a point i on the boundary of the bounding box $\delta^{\cdot}_{ij}$ are known for any $j = 1 \ldots 6$. Thus, there remain only six cases to consider.

Figure 2: The three possibilities for optimal tours

	$n=6$		$n=7$		$n=8$	
	Violations per 10^9	Max violation (%)	Violations per 10^9	Max violation (%)	Violations per 10^9	Max violation (%)
L_1	8804.5	6.953	3802.0	8.347	1399.5	8.213
$L_{4/3}$	3988.0	5.679	1707.5	6.093	768.5	7.016
L_2	1478.5	3.323	676.5	3.553	344.5	4.001
L_4	1781.5	5.098	690.0	4.542	280.0	5.837
L_∞	3024.0	5.790	1173.5	6.302	455.0	6.886
	$n=9$		$n=10$		$n=11$	
L_2	160	4.803	74	5.563	48.5	4.603

Table 2: The summary of the experiments.

Moreover, in each of the 15 cases above, the number of possible optimal tours on $I \setminus i$, $i = 1 \dots 6$, may be reduced by eliminating the tours that do not respect the order of the points on the convex hull of $I \setminus i$.

Remark. *It is possible to use the observations, similar to the ones made in the previous remark, also for the case $n = 7$. It may actually lead to resolving the seven-point case as well.*

Remark. *Any instance of geometric TSP in the L_1 metric becomes an instance of TSP in the L_∞ metric after the rotation and scaling. Thus, our results for the L_1 metric give similar results for the geometric TSP in the L_∞ metric automatically. Nevertheless, we ran a separate program for finding an instance with the maximum TSP ratio in the L_∞ case. Not surprisingly, the result was the rotated and scaled instance of the L_1 case.*

4 Experimental Study of the TSP Ratio for Small Instances

To see how often one might expect to have the TSP ratio exceed 1 for a random TSP instance, we obtained the ratio for the geometric TSP instances on $n = 6 \dots 11$ points in L_p metric for each $p \in \{1, 4/3, 2, 4, \infty\}$; the points were randomly uniformly distributed in the 100×100 square. The number of instances solved for each n and p was $2 \cdot 10^9$. For the instances of sizes 9, 10 and 11, for computational reasons, only the L_2 metric was examined. We present the results in Table 2.

For $n = 6, 7$ and 8, the maximum number of violations of the inequality $f(I) \leq 1$ occurred in the L_1 metric, the minimum number — in the L_2 or the L_4 metric. The maximum observed violation follows a similar trend.

For the L_2 metric, our results suggest an exponential decrease in the number of violations as n increases. The maximum violation though, does not monotonically decrease with n (cf. Example 2.6).

In Fig. 3 we present the number of violations in the L_2 metric as a function of the violation percentage.

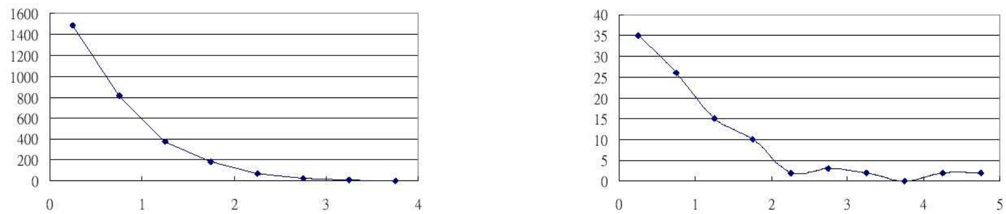

Figure 3: The distribution of the number violations in the L_2 metric. Left: $n = 6$, right: $n = 11$.

5 Conclusion

This paper introduces the notion of the TSP ratio and studies the following question: what is the range of the values that the ratio can take? It is an interesting question for combinatorial optimization, operations research, and logistics practice where individual costs have to be determined in the context of a cost for a collective.

We present a series of general upper and lower bounds on the TSP ratio. The main focus of our work is on six-node instances, for which, by solving a mixed integer program, we provide tight upper bounds on the ratio.

We computed the TSP ratio for a large number of random TSP instances on small number of points and observed that the TSP ratio exceeded 1 very rarely.

The main remaining open problem is bounding the TSP ratio for geometric TSP instances in general. Currently, the best upper bound we have is 2 — the upper bound on the TSP ratio of any metric TSP instance.

Acknowledgements

We would like thank the anonymous referees for their helpful comments. We are also indebted to Joseph Mitchell for providing the results obtained by him and his colleagues, and for the help with the presentation of the paper.

References

[1] M. Dror, Y. Lee and J. Orlin. The Metric TSP and the Sum of its Marginal Values. *Abstracts of the Fall Workshop on Comp. Geom.*, Nov. 2004, MIT, Boston.

[2] G. Gutin and A. Punnen. *Traveling Salesman Problem and Its Variations.* 2002.

[3] M. Dror and J. Ferland. Cost allocation in combinatorial problems - the TSP case, 1985. Centre de recherche sur les transports - Publication No. 435, University of Montreal.

[4] K. Juhnke and J. S. B. Mitchell. Personal communication.

[5] J. Bondy and U. Murty. *Graph Theory and Applications.* American Elsevier, New York, 1976.

[6] C. Thomassen. Hypohamiltonian and hypotraceable graphs. *Discr. Math*, (9):91–96, 1974.