A Priori Solution of a Traveling Salesman Problem in Which a Random Subset of the Customers are Visited
Author(s): Patrick Jaillet
Source: *Operations Research*, Vol. 36, No. 6 (Nov. - Dec., 1988), pp. 929-936
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/171123
Accessed: 06/06/2011 09:07

A PRIORI SOLUTION OF A TRAVELING SALESMAN PROBLEM IN WHICH A RANDOM SUBSET OF THE CUSTOMERS ARE VISITED

PATRICK JAILLET

Ecole Nationale des Ponts et Chaussees, Paris, France

(Received January 1986; revisions received December 1986, September 1987; accepted November 1987)

Consider a problem of routing through a set of n points. On any given instance of the problem, only a subset consisting of k out of n points ($0 \leq k \leq n$) has to be visited, with the number k random with known probability distribution. We wish to find a priori a tour through all n points. On any given instance, the k points present will then be visited in the *same order* as they appear in the a priori tour. The problem of finding such a tour of minimum length in the expected value sense is defined as a Probabilistic Traveling Salesman Problem (PTSP). What distinguishes one PTSP from another is the probability distribution (or more generally, the probability "law") that specifies the number k and the identity of the points that need to be visited on any given instance of the problem. After motivating the problem by applications, we first derive closed form expressions for computing efficiently the expected length of any given tour under very general probabilistic assumptions. We then provide, in a unified way, an analysis of these expressions and derive several interesting properties of the problem.

The Traveling Salesman Problem (TSP) is perhaps the most intensively investigated of all combinatorial optimization problems (Bellmore and Nemhauser 1968, Klee 1980, Parker and Rardin 1983, Lawler, Lenstra and Rinnooy Kan 1985). The effort spent on this problem is a reflection partly on the fact that this problem is an essential component of many other routing problems and that it also has other numerous, and occasionally surprising, applications (see, for example, Lenstra and Rinnooy Kan 1975).

More generally, the scholarly literature devoted to routing problems in a deterministic context has been growing rapidly over the last several years (see Bodin, Golden, Assad and Ball 1983 with about 700 references!). By a deterministic context, we mean situations in which the number of "customers," their locations and the size of their demands are known with certainty before the routes are designed. One can identify, however, a practically endless variety of problems in which one or more of these parameters are random variables, i.e., subject to uncertainty in accordance with some probability distribution. In fact, these problems, when specified in a probabilistic context, are often more applicable than their deterministic counterparts.

Let us consider the following situation: assume a company wants to design a tour through n customers and desires to minimize only the routing cost: it is then legitimate to solve the corresponding TSP as if *all* the customers must actually be visited every day. Assume, however, that this tour is to be used for a given prolonged period of time (more than one day) and that for this time horizon, the set of customers to be visited on a daily basis varies. Moreover, assume the company *cannot* reoptimize (due to lack of advance information) or simply *does not desire* to reoptimize the route every day because it is either too expensive to do so or because it prefers regularity of service. The vehicle will then follow a predesigned tour every day and, on any given day, will simply skip the missing customers from the original tour. As an illustration of such a situation, one can think of the actual case of a postman who delivers mail according to a fixed assigned route. On any particular day, upon delivery at a given location, he checks what address has to be visited next on the regular route and proceeds accordingly. The problem is not a TSP anymore since the tour must be a "good" one (small routing cost) when all customers are present, but it must also remain "good" when some customers are skipped from the original set. We have no guarantee that an optimal TSP tour through all the *potential* points has this desirable property.

This simple observation suggests the formulation and analysis of the following generic problem: Consider a problem of routing through a set of n points. On any given instance of the problem, only a subset consisting of k out of the n points ($0 \leq k \leq n$) must be visited, with the number k determined according to a known probability distribution. We wish to find a priori a tour through all n points. On any given

Subject classification: Networks/graphs: traveling salesman; probabilistic version.

0030-364X/88/3606-0929 $01.25

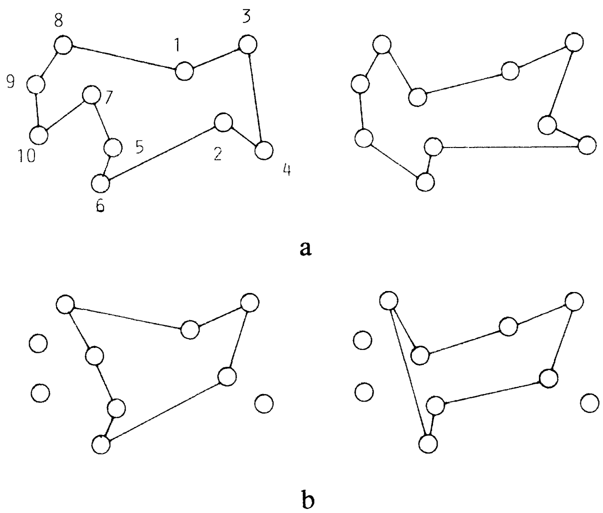

Figure 1. Simple graphical example of a PTSP. (a) Two a priori tours through the same set of 10 points. (b) The two resulting tours when points 4, 9 and 10 do not need to be visited.

instance of the problem, the k points present will then be visited *in the same order* as they appear in the a priori tour (see Figure 1 for an illustration). The problem of finding such an a priori tour, which is of minimum length in the *expected value* sense, is defined as a Probabilistic Traveling Salesman Problem (PTSP).

The previous example of the postman can be modeled by considering the simplest possible version of this general framework. By calling p the probability that any particular address will require a visit on any given day (assuming independence between addresses and an equal p for all addresses), the number of addresses requiring a visit is a binomial random variable.

We have introduced and motivated the PTSP through examples in the context of routing problems (involving physical traveling). The generic model, as stated, can be of interest in any situation in which an a priori sequence of entities has to be found for which the relative order has to be preserved even when some of the entities are absent.

Let us mention some examples. In the area of job sequencing, consider the problem of loading n jobs on a machine in which a changeover cost is involved whenever a new job is loaded. With any given ordering of n jobs on the machine, we can then associate the sum of the changeover costs. (The problem of loading the jobs in order to minimize this total cost can be formulated as a TSP; see Gilmore and Gomory 1964.) A given ordering of the n jobs may also impose specific long-term requirements such as a set of tasks to be performed before and after the processing of the jobs on the machine (think of the organization of a firm preceding and following the processing of the jobs). Very often, these requirements are difficult to modify on a daily basis so that, if on a given day some jobs do not need to be processed, we nevertheless do not modify the relative ordering previously found. The PTSP is relevant in modeling such situations as well. Finally, in the area of warehouse operations, retrieval is commonly sequenced by simply visiting storage bins according to their bin number (see Bartholdi and Platzman 1988 for a discussion on this subject, as well as the consideration of spacefilling curves for obtaining "presequences"). The number of the bins can again be modeled as a PTSP problem.

After specifying the notation and the probabilistic assumptions to be used throughout the paper (Section 1), we will present several results obtained on the PTSP. In Section 2, we present the derivation of closed form expressions for computing efficiently the expected length of any given tour. In Section 3, we show through a simple example that the optimal TSP tour can be a very poor solution to the corresponding PTSP problem. In Section 4, we investigate the links between the TSP and PTSP problems and give properties of optimum PTSP tours. In the conclusion (Section 5), we mention some related results.

1. Notation and Assumptions

Throughout this paper $G = (N, A, d)$ denotes a complete, loopless, directed weighted graph where:

N is the node set of cardinality $|N|$,

A is the set of arcs joining the nodes of $N(|A| = |N|(|N| - 1))$,

d is a function: $A \rightarrow \mathbb{R}^+$; ($d(i, j)$ represents the weight of (i, j), i.e., the direct cost or distance from node i to node j).

t represents a Hamiltonian circuit (*tour*) of G; by reindexing the nodes in their order of appearance along t, we write t as

$$t = (1, 2, 3, 4, \ldots, |N|, 1).$$

The set of nodes N is partitioned into two subsets N_1 and N_2 ($N_1 \cup N_2 = N$, $N_1 \cap N_2 = \varnothing$).

N_1 is the set of nodes that will always require a visit (be present) for each instance of the problem; $|N_1| = m$.

N_2 is the set of nodes that will *not* always require a visit for each instance of the problem; $|N_2| = n$.

We use the terms black nodes and white nodes for the elements of N_1 and N_2, respectively.

We assume that given a probability distribution P on Ω, the power set of N_2; an outcome ω defines the subset of white nodes that require a visit. We *restrict* P to be such that all outcomes of same cardinality have the same probability of occurring

$$(\text{for all } \omega_1 \in \Omega \text{ for all } \omega_2 \in \Omega,$$
$$|\omega_1| = |\omega_2| \Rightarrow P(\{\omega_1\}) = P(\{\omega_2\})). \quad (\#)$$

If K is the random variable that represents the number of white nodes that require a visit we have

$$P(K = |\omega|) = \binom{n}{|\omega|} P(\{\omega\}).$$

Hence, our probabilistic models can be specified equivalently by giving the probability P or the probability distribution of K. Note also that the restriction imposed on P implies that, given $K = k$, the k nodes are taken uniformly at random among the set of n nodes; any probability P satisfying (#) will then be said to be *node-invariant* (NI).

One important specific example (hereafter named P_1) is

$$P_1(\{\omega\}) = p^k(1-p)^{n-k}, \quad \text{with } k = |\omega| \qquad (\#\#)$$

which corresponds to the case for which each white node has a probability p of being present, independently of the others; we then speak informally of a Bernouilli process with parameter p.

In Sections 2.1 and 3 we assume $P = P_1$; *in all* other sections we assume a general node-invariant probability.

For a given a priori tour t, the length L_t covered in traversing the set of nodes actually present on each instance of the problem is a random variable. The general PTSP can then be stated as follows.

Problem PTSP

Given $G = (N, A, d, P)$ find an a priori tour t of minimum expected length, $E[L_t]$.

In Section 4, two specific tours play an important role:

$t1$: an optimal TSP tour through all $m + n(= |N|)$ nodes

tp: an optimal PTSP tour.

2. The Expected Length of a Given Tour *t*

Let us consider a PTSP problem defined as $G = (N, A, d, P)$. For a given tour t, L_t is a random variable that can have up to 2^n different values. By considering all cases, its expected value would then be obtained in $O((n + m)2^n)$ additions in the worst case, and this is not satisfactory. On the other hand, $E[L_t]$ can clearly be expressed as the sum over all arcs (i, j) of $P($"(i, j) appears in a subtour") times $d(i, j)$. So if $P($"(i, j) appears in a subtour") can be computed efficiently, say in $O((n + m)^k)$, we would obtain $E[L_t]$ in $O((n + m)^{k+2})$.

The purpose of this section is to show that, with a node-invariant P, we can obtain these probabilities efficiently, but, more importantly, we can express $E[L_t]$ in terms of a set of well defined quantities, whose analysis (Section 4.1) proves to be fundamental in deriving properties of optimum PTSP tours (Section 4.2).

2.1. Case of a Bernouilli Process

In this subsection we assume that $P = P_1$, i.e., each white node is present with a probability p, independently of the others.

Theorem 1. *Given a graph $G = (N, A, d, P_1)$ where $|N_1| = m$, $|N_2| = n$, and P_1 corresponds to a Bernouilli process with parameter p, the expected length $E[L_t]$ of a tour $t = (1, 2, \ldots, n + m, 1)$ is*

$$E[L_t] = p^2\left[\sum_{r=0}^{n-2} (1-p)^r L_{m,t}^{(r)}\right] + p(1-p)^{n-1}L_{m,t}^{(n-1)} + (1-p)^n L_{m,t}^{(n)} \qquad (1)$$

where

(i) $L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}(j, \overline{j+r+1})$ *for all* $r \in [0 \,..\, n-1]$

with

$\bar{k} = (k-1)$ *modulo* $(n+m) + 1$ *for* $k \geq 1$

$d_{m,t}(j, \overline{j+r+1}) = d(j, \overline{j+r+1})$ *whenever the nodes* $\overline{j+1}, \overline{j+2}, \ldots, \overline{j+r}$ *are all white nodes.*

$d_{m,t}(j, \overline{j+r+1}) = \sum_{c=0}^{s} d(k_c, k_{c+1})$ where $k_0 \equiv j$ $k_{s+1} \equiv \overline{j+1+r}$ *and* $(k_1, k_2, \ldots, k_s)$ *is the sequence of black nodes drawn from* $(\overline{j+1}, \ldots, \overline{j+r})$.

(ii) $L_{m,t}^{(n)}$ *is the length of the tour t through the m black nodes (i.e., when no white nodes are present).*

Proof. On any given instance of the problem, the arc $(j, \overline{j+r+1})$ is in the resulting tour if and only if

- the nodes j and $\overline{j+r+1}$ are present,
- the nodes $\overline{j+1}, \overline{j+2}, \ldots, \overline{j+r}$ are absent and thus skipped.

Four cases have to be considered to evaluate the probability of presence of this arc:

1. if at least one node among $\overline{j+1}, .., \overline{j+r}$ is a black node, the probability of presence is 0, otherwise;
2. if nodes j and $\overline{j+r+1}$ are white nodes, the probability is $p^2(1-p)^r$;
3. if nodes j or $\overline{j+r+1}$ is a black node, the probability is $p(1-p)^r$;
4. if nodes j and $\overline{j+r+1}$ are black nodes, the probability is $(1-p)^r$.

For each case, the probability of presence does not depend on j but only on r, so that one can regroup arcs $(j, \overline{j+r+1})$ that belong to the same cases for a given r. The $L_{m,t}^{(r)}$'s represent one way of regrouping arcs based on this idea; note, however, that cases 2, 3, and 4 do not correspond to each of the three terms in the expression on a one-to-one basis (see the Appendix for details).

For an intuitive understanding of these quantities, it is worth mentioning that when $m = 0$, $L_{0,t}^{(r)}$ is the sum of n elements, each representing the distance from node j to its $(r+1)$th "successor" along the tour t. For $m > 0$, the definitions can be kept similar, once we introduce a "transformed" distance $d_{m,t}$ which reflects the possibility of having black nodes (hence, never skipped) between node j and node $\overline{j+r+1}$ along the tour t.

2.2. Generalizations

Theorem 1 assumes that $P = P_1$ or, equivalently, that K is a binomial random variable. Under the node-invariant property, this result can be generalized to the case of a general discrete probability distribution for K. Expression 1 is still valid under this general setting if one replaces

$$p^2(1-p)^r \quad \text{by} \sum_{k=r}^{n-2} \left[\binom{n-2-r}{k-r} \Big/ \binom{n}{k} \right] P(K = n-k) \quad \text{for all } r \in [0 \,..\, n-2]$$

$$p(1-p)^{n-1} \quad \text{by } P(K=1)/n \tag{2}$$

$$(1-p)^n \quad \text{by } P(K=0).$$

Indeed, let us consider case 2 of the previous proof. Now *given* $K = n - k$ (i.e., k nodes do not require a visit), the probability of presence of arc $(j, \overline{j+r+1})$ is

$$\binom{n-2-r}{k-r} \Big/ \binom{n}{k} \quad \text{if } r \leq k$$

$$0 \quad \text{otherwise,}$$

($k - r$ nodes "not present" have to be chosen among the $n - 2 - r$ yet unchosen nodes). Hence, the unconditional probability of the event is given by the first line of (2). The other cases are treated in a similar fashion, and by regrouping arcs as described in Theorem 1, we obtain the desired result.

Comments

1. One can also apply the same technique to compute the expected length of a path instead of a tour. The closed form expressions remain identical after the definitions of the L_m's are slightly modified. For more detailed results on this matter, as well as on the PTSP in the special cases for $m = 0$ and $m = 1$, the reader is referred to Jaillet (1985).
2. One can define a broader class of probability distributions P for which $E[L_t]$ can efficiently be computed. The general property is that for any partition of N_2 into subsets A, B, C, one can easily compute the probability that A is in the subtour and B is not where C does not matter. Of course, with such a general P, we are no longer able to express $E[L_t]$ in terms of the $L_{m,t}$ so that the results of Section 4 do not generally follow.
3. The complexity of computing $E[L_t]$ (for a node-invariant P and by using (1) and (2)) is $O(m + n^2)$.

3. Using the TSP Solution to Solve a PTSP Counterexample

Let us present an example showing that an optimal TSP tour is not necessarily a good solution to the corresponding PTSP problem. In this example, the graph G contains 24 white nodes (and no black node) that are positioned at the vertices of two concentric 12-gons as shown in Figure 2a. We assume $P = P_1$ (Bernouilli process with parameter p). In Figure 2b, two tours have been designed through this set of nodes: tour a is an optimal TSP tour, tour b is an alternative tour (see Jaillet 1985 for the numerical derivations). One can then show that for a probability of presence of 0.5, the expected length of tour a is 31% greater than the expected length of tour b. This numerical example raises the following question: in general, how well would a TSP tour do as a solution to a PTSP problem? In the next section we address, among other topics, this particular question.

4. Properties and Characteristics of Optimal Solutions to PTSP's

This section shows that the PTSP introduces many features that are different from those of its famous

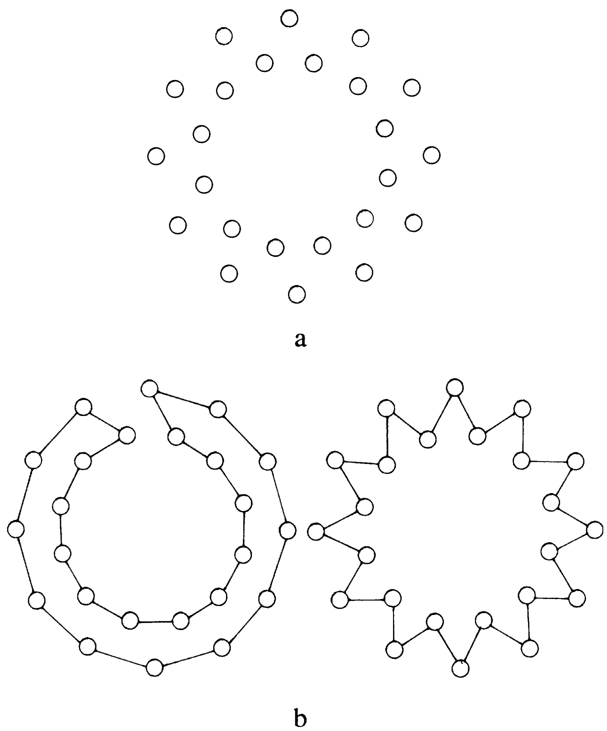

Figure 2. Graph and tours for the numerical example. (a) A 24-node graph. (b) Two tours of the 24-node graph.

special case, the TSP. The TSP is a special case of the PTSP in which all the nodes are black; it is then natural to investigate the possible links between the two problems. We will be concerned with two different, but related, issues in the next two theorems: first we examine the question of how far an optimal TSP tour can be from optimality for a PTSP problem (Theorem 2); then we determine conditions, if any, under which the PTSP problem is solved by the optimal TSP tour (Theorem 3).

4.1. Analysis of the Closed Form Expression

Most of the properties obtained in this section are derived from a detailed analysis of the generalized closed form expression of Theorem 1; in what follows, we give the main results of this analysis in the form of two lemmas (see Jaillet 1985 for the proofs) for the case of a general distribution for K.

The expression derived in Section 2 can be written as follows in a general weight-form notation

$$E[L_t] = \sum_{r=0}^{n} \alpha_r L_{m,t}^{(r)}$$

where

$$\alpha_r = \sum_{k=r}^{n-2} \left[\binom{n-2-r}{k-r} \Big/ \binom{n}{k} \right] P(K = n-k) \qquad \text{for any } r \in [0 \,..\, n-2]$$

$$\alpha_{n-1} = P(K=1)/n$$

$$\alpha_n = P(K=0).$$

$E[L_t]$ is thus composed of two families of elements: the α_r and the $L_{m,t}^{(r)}$.

We will always assume that $n \geqslant 1$; $n + m \geqslant 3$. (Note that for $n = 0$ the PTSP reduces to the traditional TSP and for $m + n \leqslant 2$ we have a unique tour; both cases are obviously of little interest for our purposes.)

a. Properties of the Weights α_r

The properties of the weights α_r are summarized by the following lemma, which is valid for any tour t.

Lemma 1. *Given any choice for the discrete probability distribution of K we have:*

(i) $\alpha_r \leqslant \alpha_{r'}$ *for* $0 \leqslant r' \leqslant r \leqslant n-2$
(ii) $\sum_{r=0}^{n-1} \alpha_r = E[K]/n$
(iii) $\sum_{r=0}^{n-1} (r+1)\alpha_r + \alpha_n = 1.$

b. Properties of $L_{m,t}^{(r)}$

They are summarized in the following lemma.

Lemma 2. *Given a graph G with n white nodes, m black nodes, and given a tour t, then:*

a. *If d is any function:*
   (i) $\sum_{r=0}^{n} L_{0,t}^{(r)}$, $L_{1,t}^{(n-1)}$, $L_{2,t}^{(n)}$ *are tour-independent.*
   (ii) $L_{0,t}^{(r)} \geqslant L_{0,t_1}^{(0)}$ *for any r such that G.C.D*$(n, r+1) = 1$.
b. *If d is symmetric:*
   (i) $L_{0,t}^{(n-2-r)} = L_{0,t}^{(r)}$ *for any* $r \in [0 \,..\, n-2]$ (*Note: this is not true for* $L_{m,t}^{(r)}$ *when* $m \geqslant 1$.)
   (ii) $\sum_{r=0}^{n} L_{1,t}^{(r)}$, $L_{2,t}^{(n-1)}$, *and* $L_{3,t}^{(n)}$ *become tour-independent.*
c. *If d satisfies the triangular inequality:*
   (i) $L_{m,t}^{(r)} \geqslant L_{m,t_1}^{(0)}$ *for any* $m \geqslant 1$ *for any* $r \in [0 \,..\, n-1]$
   (ii) $L_{m,t}^{(r)} \leqslant L_{m,t}^{(r1)} + L_{m,t}^{(r2)}$ *for any* $m \geqslant 0$ *for any* $r \in [1 \,..\, n-1]$ *for any* $r1, r2{:}\ r1 + r2 = r - 1$.

4.2. Some Properties of Optimum PTSP Tours

The material covered in Section 4.1 turns out to be instrumental in deriving several interesting results concerning the PTSP. We present some of these here under the assumption of a general node-invariant probability P.

Theorem 2. *Given a graph G with n white nodes, m black nodes, a distance function d that satisfies the triangular inequality, an optimal PTSP tour tp on G, and an optimal TSP tour t1 on G, we have*

$$(E[L_{t1}] - E[L_{tp}])/E[L_{tp}]$$
$$\leq (1 - E[K]/n)/(E[K]/n)$$

for any $m \geq 1$ for any n (for $m = 0$ for any n prime).

Proof. For $m \geq 1$: From Lemma 2 (ci) and Lemma 1 (ii) we have

$$E[L_{tp}] = \sum_{r=0}^{n} \alpha_r L_{m,tp}^{(r)} \geq L_{m,t1}^{(0)}(E[K]/n) + \alpha_n L_{m,tp}^{(n)}. \quad (3)$$

From Lemma 2 (cii) and Lemma 1 (iii) we have

$$E[L_{t1}] = \sum_{r=0}^{n} \alpha_r L_{m,t1}^{(r)} \leq L_{m,t1}^{(0)}(1 - \alpha_n) + \alpha_n L_{m,t1}^{(n)}. \quad (4)$$

Combining (3) and (4) we obtain

$$E[L_{t1}] - E[L_{tp}] \leq L_{m,t1}^{(0)}[1 - E[K]/n]. \quad (5)$$

Dividing (5) by (3) the theorem is proved for $m \geq 1$.

For $m = 0$ and n prime, the proof is similar. (We conjecture but have not been able to prove that the theorem also holds for all n if $m = 0$.)

Comments

1. When K is a binomial random variable with parameter p, Theorem 4 gives $(E[L_{t1}] - E[L_{tp}])/E[L_{tp}] \leq (1 - p)/p$.

2. For $m = 1$, $L_{1,t}^{(n)} = 0$ so that (4) and (5) give a slightly better bound, namely

$$(1 - E[K]/n - \alpha_n)/(E[K]/n).$$

3. When $E[K]/n$ approaches zero, the sharpness of this bound is questionable. Note that when $E[K]/n = 1$, the bound gives the correct answer. In fact, one can show (using a generalization of the star-shaped example of Section 3) that the TSP can indeed be *arbitrarily* bad under the conditions of Theorem 2.

Theorem 3. *Given a problem instance with n white nodes and m black nodes, an optimum TSP tour solves the PTSP optimally for any underlying graph G = (N, A, d) and for any probability distribution of K if and only if:*

(i) *d is symmetric and $m + n \leq 4$; if $m = 0$ this is also true for $n = 5$.*
(ii) *d is not symmetric and $m + n \leq 3$.*

Proof. If: using Lemmas 1 and 2 one can show that $E[L_t]$ is, in each case, a linear function of $L_{m,t}^{(0)}$ with positive slope so that

$$\min_t E[L_t] \equiv \min_t L_{m,t}^{(0)}.$$

Only if: to prove the "only if" proposition it suffices to construct a counterexample (see Jaillet 1985) for each of the following cases:

1. d symmetric, $m = 1, n = 4$, K binomial,
2. d symmetric, $m = 0, n = 6$, K binomial,
3. d not symmetric, $m = 1, n = 3$, K binomial.

To conclude the investigation of the relationship between TSP and PTSP let us give a last result.

Lemma 3. *Let G = (N, A, d) be a given graph lying in R^2 with d the Euclidean metric. Let $\mathscr{C}(N)$ be the set of points of N that belong to the convex hull of $N(\mathscr{C}(N) \subset N)$:*

(i) *if $\mathscr{C}(N) = N$, then the solutions of the TSP and PTSP problems are identical for any probability distribution of K.*
(ii) *if $\mathscr{C}(N) \neq N$, one can construct instances for which the optimal PTSP tour intersects itself.*

Proof

(i) Based on the following well known property for the Euclidean TSP: the order in which the points on the convex hull appear in an optimal TSP tour must be the same as the order in which these points appear on the convex hull (see, for example, Larson and Odoni 1981).
(ii) By construction (Figure 3), the exact calculations are given in Jaillet (1985).

5. Conclusion

In this paper, we have introduced the PTSP, a problem that provides a conceptual model for many practical situations likely to be encountered in various forms in several application areas. We note that in addition to the specific examples mentioned in the Introduction, the PTSP methodology could be used in other areas, such as facility location or preliminary planning for routing problems (see Jaillet 1985).

Following the formulation of the PTSP, we then presented closed form expressions to obtain efficiently (i.e., in polynomial time of low order) the expected length of tours under various probabilistic assumptions. More important than just the efficiency in computing $E[L_t]$ is the fact that these closed form expressions express $E[L_t]$ in terms of a set of well defined quantities (namely, the $L_{m,t}$). Understanding

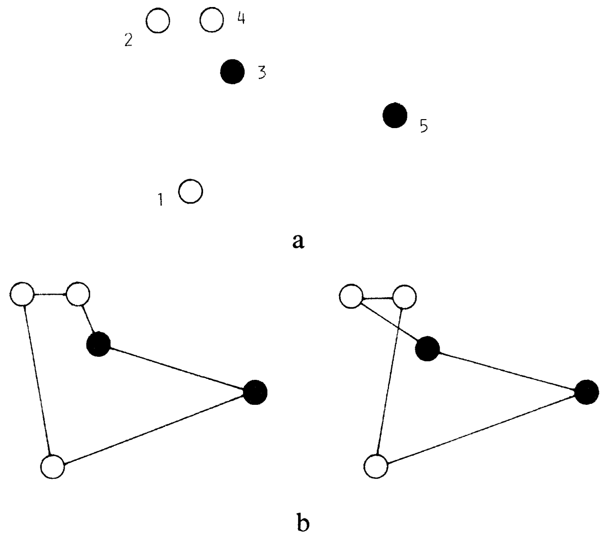

Figure 3. Intersection of the optimal PTSP tour. (a) The set of five points. (b) The optimal PTSP tours. Left side: optimal tour 1 when $p > 0.75$; right side: optimal tour 2 when $p < 0.25$.

the properties of these quantities proved to be fundamental in comparing the TSP and PTSP.

The analysis presented in this paper implies that under specific conditions the TSP solution can serve as a good approximation for the PTSP problem. In general, however, the problem presents sufficiently different features from its special case to necessitate devising entirely new solution procedures. In Jaillet (1985) we propose a branch-and-bound scheme and several heuristic methods, but this remains of an introductory nature and much has yet to be discovered and tested in this difficult area.

A forthcoming paper will present a comprehensive analysis of the PTSP in the plane that eventually leads to interesting asymptotic results. These latter results, obtained in the limit (as the number of points m and/or n grows to infinity) are in the same spirit as those of Beardwood, Halton and Hammersley (1959) for the TSP. Let us also mention that we obtained some results on probabilistic versions of other well known problems such as the vehicle routing problem (Jaillet 1987a) and the shortest path problem (Jaillet 1987b).

Appendix

End of the Proof of Theorem 1

The purpose of this appendix is to fill in the gap between (1) and the three cases considered in the proof of Theorem 1 (see Section 2.1).

First of all, it is important to note that we have arcs of

$$\begin{aligned} &\text{case 2} \quad \text{iff } r \leq n-2, \\ &\text{case 3} \quad \text{iff } r \leq n-1, \\ &\text{case 4} \quad \text{iff } r \leq n, \end{aligned} \tag{A.1}$$

since we have a total of n white nodes.

Now, by carefully following the definition of the $L_{m,t}$ one can see that arcs of

$$\begin{aligned} &\text{case 2} \quad \text{are involved in } L_{m,t}^{(r)} \text{ only,} \\ &\text{case 3} \quad \text{are involved in } L_{m,t}^{(k)} \\ &\qquad\quad \text{for } k \in [r \,..\, n-1], \\ &\text{case 4} \quad \text{are involved } k+1-r \text{ times in } L_{m,t}^{(k)} \\ &\qquad\quad \text{for } k \in [r \,..\, n-1] \text{ and once in } L_{m,t}^{(n)}. \end{aligned} \tag{A.2}$$

So (A.2) implies that the weight of the contribution of each arc in (1) is, respectively, for

case 2: $p^2(1-p)^r$,

case 3: $p^2\left[\sum_{k=r}^{n-2} (1-p)^k\right] + p(1-p)^{n-1} = p(1-p)^r$,

case 4: $p^2\left[\sum_{k=r}^{n-2} (k+1-r)(1-p)^k\right]$

$$+ (n-r)p(1-p)^{n-1} + (1-p)^n = (1-p)^r,$$

and these correspond to the probabilities of presence as derived previously in Section 2.1. This fact, combined with (A.1), terminates the proof of Theorem 1.

Acknowledgment

The author is grateful to Professor Amedeo Odoni for his numerous comments on earlier versions of this paper and for supervising the Ph.D. dissertation on which much of this paper is based. Also, many thanks to an anonymous referee for useful comments for improving the presentation, for suggesting the present title, and pointing out the general property listed in Comment 2 of Section 2.2.

References

Bartholdi, J., and Platzman. 1988. Heuristics Based on Spacefilling Curves for Combinatorial Problems in Euclidean Space. *Mgmt. Sci.* **34**, 291–305.

Beardwood, J., J. Halton and J. Hammersley. 1959. The Shortest Path Through Many Points. *Proc. Camb. Phil. Soc.* **55**, 299–327.

Bellmore, M., and G. Nemhauser. 1968. The Traveling Salesman Problem: A Survey. *Opns. Res.* **16**, 538–558.

BODIN, L., B. GOLDEN, A. ASSAD AND M. BALL. 1983. Routing and Scheduling of Vehicles and Crews, the State of the Art. *Comput. Opns. Res.* **10,** 69–211.

GILMORE, P., AND R. GOMORY. 1964. Sequencing a One-State Variable Machine: A Solvable Case of the Traveling Salesman Problem, *Opns. Res.* **12,** 655–679.

JAILLET, P. 1985. The Probabilistic Traveling Salesman Problems. Technical Report 185, Operations Research Center, MIT, Cambridge, Mass.

JAILLET, P. 1987a. Stochastic Routing Problems. In *Stochastics in Combinatorial Optimization*, G. Andreatta, F. Mason, P. Serafini (eds.). World Scientific.

JAILLET, P. 1987b. On Some Probabilistic Combinatorial Optimisation Problems Defined on Graphs. In *Flow Control of Congested Networks*, A. Odoni, L. Bianco, and G. Szego, (eds.). NATO ASI Series F, **38,** Springer-Verlag, Berlin.

KLEE, V. 1980. Combinatorial Optimization: What Is the State of the Art. *Math. Opns. Res.* **5,** 1–26.

LARSON, R., AND A. ODONI. 1981. *Urban Operations Research.* Prentice Hall, Englewood Cliffs, N.J.

LAWLER, E., J. K. LENSTRA AND A. RINNOOY KAN. 1985. *The Traveling Salesman Problem*, John Wiley & Sons, New York.

LENSTRA, J., AND A. RINNOOY KAN. 1975. Some Simple Applications of the Traveling Salesman Problem, *Opns. Res. Quart.* **24,** 717–733.

PARKER, R., AND R. RARDIN. 1983. The Traveling Salesman Problem: An Update of Research. *Naval Res. Logist. Quart.* **30,** 69–96.