Yugoslav Journal of Operations Research
15 (2005), Number 1, 5-14

APPLICATIONS OF A SPECIAL POLYNOMIAL CLASS OF TSP

Jack BRIMBERG,
Royal Military College of Canada, CP17000 Stn Forces
Kingston, Ontario, Canada K7K 7B4
Ephraim KORACH
Ben-Gurion University of the Negev, Beer-Sheva, Israel
Mokhtar AMAMI
Royal Military College of Canada, CP17000 Stn Forces
Kingston, Ontario, Canada K7K 7B4

Presented at XXX Yugoslav Simposium on Operations Research
Received: January 2004 / Accepted: January 2005

Abstract: A hypothetical problem which we call a "buried treasure problem" is presented where the objective is to locate *m* objects among *N* fixed equi-spaced caches in order to minimize a measure of the risk of loss. The general problem is shown to be NP-hard. However, a sub problem may be solved as a special class of TSP in O (N log N) time. Several applications are noted.

Keywords: Partition problem, traveling salesman problem, pyramidal tours.

1. INTRODUCTION

A fictional scenario is presented which we term a "buried treasure problem". The objective is to partition and locate *m* objects of value among *N* secret caches in order to minimize a measure of the risk of loss. Under the model assumptions, the general problem is shown to be reducible to the partition problem which is *NP-complete* (e.g., Garey and Johnson (1979), p. 60-61; Simeone (1986)). However, for a specified partitioning of the set of objects, the assignment of subsets to the individual caches may be formulated as a special class of traveling salesman problem (TSP), and solved in polynomial time.

The distance matrix of the formulated TSP is shown to possess the anti-Monge property (e.g., see Burkard et al. (1995) for a definition). For a general Monge matrix, it is well known that an optimal tour may be constructed with a pyramidal sequence in $O(N^2)$ operations using dynamic programming (see Lawler, Lenstra, Rinnooy Kan and Shmoys (1990), Chapter 4; and Burkard (1997)). However, the special structure of our TSP enables us to explicitly present the optimal tour by a specified permutation of the ordered weights of the *N* subsets. If the weights are already sorted in non-increasing order, the optimal tour is constructed directly in linear time; otherwise, the algorithm requires $O(N \log N)$ initial operations to complete the sorting.

We also note that the distance matrix in our model is a symmetric product matrix for which there are known results on the optimal Hamiltonian circuit for both non-positive and nonnegative cases (e.g., see the excellent survey in Burkard et al., 1998). However, to our best knowledge, the results given here for the TSP path, although simple to derive, are new. The derivation of the optimal permutation is provided for completeness for the minimum and maximum traveling salesman path and circuit, all of which are solved explicitly from the sorted sequence of weights. Note that the minimum path and circuit are non-pyramidal, while the maximums are pyramidal (as expected by the Monge property).

In addition to presenting the "buried treasure problem" and deriving some analytical results, this paper discusses several interesting applications, and proposes some directions for future research. The only known previous application of the TSP with symmetric product matrix is found in Hallin et al. (1992), where minimizing the autocorrelation coefficient for a given time series is analyzed. These authors solve for the minimum and maximum TSP circuits; however, their proof of optimality is much more complicated than the proof we give here. The maximum circuit is given previously in Aizenshtat and Kravchuk (1968), again with a different method of proof.

2. THE PROBLEM

Consider the following problem. We have *m* objects of value which we would like to hide among $N \geq 2$ fixed equi-spaced caches along a road. The value or worth of object *j* is given by a known constant, $y_j > 0$, $j = 1,..., m$. We have been tipped off by friendly sources that a raid will be conducted by a gang of thieves tonight over a randomly selected continuous segment of the road. Furthermore, we anticipate that due to a time limit, the length ($2l$) of this segment will be greater than the distance (a) between two adjacent caches, but less than twice this distance ($a < 2l < 2a$).

The problem confronting us is how to assign the objects among the caches in order to avoid "as much as possible" any losses incurred by the raid. Since we have no information concerning the location of the search interval along the road, we must assume that any segment of length $2l$ is equally likely. Let *X* denote the centre of the segment, and consider the section of road of length, $L = a(N - 1) + 2l$, containing the caches shown in Figure 1. We have assumed therefore that the probability density function of *X* is given by:

$$f(x) = c, \quad 0 \leq x \leq L, \tag{1}$$

where $c \le 1/L$ is constant. Let the total value of the objects stored in cache *i* be denoted by v_i, $i = 1, ..., N$. Then, given that *X* falls in the interval [0, *L*], at least one cache and possibly two will be discovered. For the position of *X* shown in Figure 1, the loss incurred would be $v_3 + v_4$.

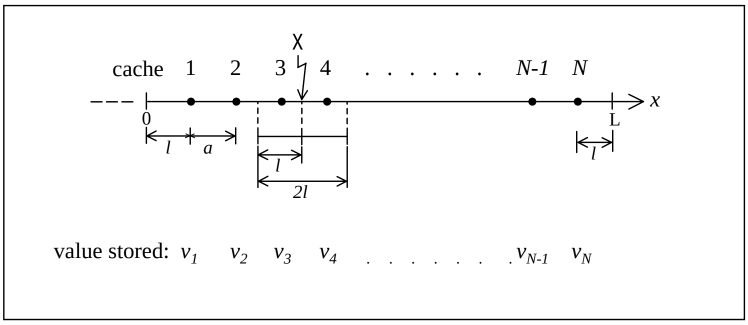

Figure 1: Illustration of the problem

Let *A* be the total value found in the search, referred to as the loss function. According to our rules,

$$A = \sum_{i=1}^{N} v_i d(|X - p_i|), \tag{2}$$

where $p_i = l + (i-1)a$ is the coordinate of cache *i*, *i* = 1 ,..., *N*, and

$$d(|X - p_i|) = \begin{cases} 1, if \ |X - p_i| \le l \\ 0, if \ |X - p_i| > l \end{cases}$$

The expected loss is readily calculated:

$$E[A] = \sum_{i=1}^{N} v_i E\left[d\left(|X - p_i|\right)\right]; \tag{3}$$

but

$$E\left[d\left(|X - p_i|\right)\right] = \int_{p_i - l}^{p_i + l} f(x)dx = 2lc; \tag{4}$$

hence

$$E[A] = 2lcW, \tag{5}$$

where

$$W = \sum_{i=1}^{N} v_i = \sum_{j=1}^{m} y_j. \tag{6}$$

Thus, the expected loss from the confiscation of objects by the raiders has a constant value no matter what we do. Whether all the objects are placed in a single cache, or dispersed among them, has no effect on this criterion. What then is the advantage of using several caches? If all the goods are hidden in only one of them, there will be a greater probability that the raiders will come up empty-handed. But if their search interval contains the selected cache, we lose everything. It follows intuitively that dispersing the objects among different caches will reduce the variability of the end result (which would be the preferred alternative for the risk-averse individual or scenario). The objective here will be to verify this intuition in a formal way. Let us therefore first evaluate the variance of A:

$$Var[A] = E\left[A^2\right] - \left(E[A]\right)^2; \tag{7}$$

after some steps, we obtain

$$Var[A] = 2c\left[l\sum_{i=1}^{N} v_i^2 + (2l-a)\sum_{i=1}^{N-1} v_i v_{i+1}\right] - 4c^2 l^2 W^2. \tag{8}$$

Thus, the problem of minimizing the variance is equivalent to assigning objects to caches in order to:

$$\text{minimize } B_1 = l\sum_{i=1}^{N} v_i^2 + (2l-a)\sum_{i=1}^{N-1} v_i v_{i+1} \tag{P1}$$

In turn, this problem may be viewed as consisting of two steps:

Step 1. {Partitioning} Determine a suitable partition of the m objects into N subsets (one per cache), $S_1, \ldots, S_N$, some of which may be empty. The value q_t of S_t would simply be the sum of the values of the objects contained in this subset:

$$q_t = \sum_{j \in S_t} y_j, t = 1, \ldots, N. \tag{9}$$

Step 2. {Location} Assign each S_t to the proper cache. This gives a sequence $S_{[i]}$, i = 1,..., N, where subset $S_{[i]}$ is assigned to cache i, and $v_i = q_{[i]}$, the value of $S_{[i]}$.

It is important to note that once a partition in Step 1 is determined, the first summation in (P1) becomes a constant. Thus, the problem in Step 2 requires assigning the subsets to the caches in order to:

$$\text{minimize } B_2 = \sum_{i=1}^{N-1} v_i v_{i+1}. \tag{P2}$$

That the cross-product terms only include adjacent pairs relies on the original assumption that the search interval contains at most two adjacent caches; thus

$$E\left[d\left(\left|X - p_i\right|\right) d\left(\left|X - p_j\right|\right)\right] = 0$$

if p_i and p_j are nonadjacent vertices.

This paper will concentrate mainly on the location problem represented by (P2). Thus, there are N groups (or subsets) of objects, $S_1,\ldots, S_N$, with corresponding non-negative values $q_1,\ldots, q_N$ ($q_t = 0$ if S_t is empty) which must be assigned one to each of N caches in order to minimize B_2. Admittedly, the burying of treasures among caches in order to minimize the variance of a loss function under the ideal conditions described above is a purely fabricated problem; yet, some practical applications may come to mind. In fact, the cross-product terms $v_i v_j$ may measure interactions in a completely different context.

For example, suppose N people will be seated around a table at a social event or business meeting such as a brainstorming session. The parameter q_i measures the tendency of the i^{th} person to converse with his neighbours, and we assume that the product $q_i q_j$ estimates the level of conversation or rate of exchange of ideas between i and j if they are seated next to each other. Alternatively, we may want to construct a circuit to connect N centres (e.g., computers). Centre i has q_i users, and the product $q_i q_j$ estimates the interaction between i and j if these two centres are linked directly to each other. It may make sense in this context to maximize the sum of interactions of adjacent centres. As another example, consider N components that must be placed at equal distances along a line. As a first approximation, only the interactions (e.g., corrosive, gravitational, magnetic) between adjacent components will be considered. These interactions are given by the product form $q_i q_j$.

The problem identified in (P2) may be formulated as a special class of traveling salesman problem. Let $G(V, E)$ denote an undirected graph, where $V = \{1,..., N\}$ is a set of N vertices representing N entities, and $E = \{(i,j)\colon i = 1,..., N\text{-}1, j = i+1,..., N\}$ is the set of edges, one between each pair of vertices. Let $r_{ij} = q_i\, q_j = r_{ji}$, $\forall\,(i,j) \in E$, where q_i is a node number at vertex i. It then follows that problem (P2) is equivalent to finding a "minimum traveling salesman path" in the graph G. Referring to our social or business event, we may alternatively want to find a "minimum or maximum traveling salesman circuit" on G depending on whether it is desirable to reduce or increase the level of conversation. Similarly, in the succeeding example, the minimum or maximum traveling salesman path on G is sought, depending on whether we wish to minimize or maximize the sum of (chemical, physical, electrical) interactions.

In the next section, the minimum and maximum traveling salesman path and circuit on G are derived. For this special class of TSP, the solutions are given by specific sequences of the node numbers (q_i) which may be obtained in polynomial $\{O(N \log N)\}$ time.

3. OPTIMAL TOURS ON G

Node i on graph G is seen to represent the entity S_i with value (or weight) given by node number q_i, $i = 1,..., N$. Furthermore, the length of any arc (i,j), $r_{ij} = q_i q_j$, measures an interaction between i and j if these two nodes occupy adjacent positions (or caches). The total length of a tour (path or circuit) on G gives a sum of interactions representative of the objective function in (P2). The applications noted above are thus equivalent to finding a tour of minimum or maximum length on G.

In order to find an optimal tour on G, the vertices are first arranged in non-increasing order of the weights q_i, as given by the following sequence:

$$w_1 \geq w_2 \geq \ldots \geq w_N \ . \tag{10}$$

Let $D = [d_{ij}]$ denote the distance matrix corresponding to this ordering of the nodes. We obtain the following preliminary result.

Property 1. *D is anti-Monge.*

Proof: For any $i < r$ and $j < s$ from the index set $\{1,..., N\}$,

$$(d_{ij} + d_{rs}) - (d_{is} + d_{rj}) = (w_iw_j + w_rw_s) - (w_iw_s + w_rw_j) = (w_i - w_r)(w_j - w_s) \geq 0. \quad \blacklozenge$$

The optimal tour is constructed directly from the ordered sequence of nodes in (10), as shown by the following results. We first consider the minimum length path and circuit on G.

Property 2. *A minimum length traveling salesman path is given by the following permutation:*

$$w_1, w_N, w_3, w_{N-2}, \ldots, w_{N-3}, w_4, w_{N-1}, w_2 \tag{11}$$

Proof: To illustrate, suppose $N > 10$, and w_5 is the first weight which does not conform to the specified sequence in (11) which we observe is being built simultaneously inwards from left and right. The current solution therefore has the following form:

$$w_1, w_N, w_3, w_{N-2}, x,\ldots, w_5, y, \ldots, w_{N-3}, w_4, w_{N-1}, w_2,$$

where y may be w_{N-3}. Consider a second solution obtained by flipping the subsequence, $x, \ldots , w_5$, in a block so that w_5 is now adjacent to w_{N-2}, the elements in the subsequence are in reverse order, and x is just to the left of the weight y. The change in path length in going from the current to the second solution is given by:

$$\Delta = (w_{N-2}\, w_5 + xy) - (w_{N-2}\, x + w_5\, y) = (w_{N-2} - y)(w_5 - x) \leq 0,$$

since $w_5 \geq x$, and $w_{N-2} \leq y$. Thus, the second solution is equivalent to or better than the current one.

A similar argument holds for any N and any first element which does not conform to the construction in (11). We conclude that a series of “flips” may be applied to any solution to obtain the permutation in (11) without increasing the path length, and hence, this sequence is a minimum path on G. ♦

Property 3. *A traveling salesman circuit of minimum length is given by the following permutation:*

$$\ldots.. , w_{N-3}, w_3 , w_{N-1}, w_1, w_N, w_2, w_{N-2}, \ldots.. \quad . \tag{12}$$

Proof: A series of flips may be used in an analogous fashion as in the path problem. ♦

It is interesting to observe in (11) and (12) that the minimum path sequence is constructed inwards from left and right using alternately the heaviest and lightest remaining elements, while the minimum circuit is constructed from the centre outwards in like fashion.

Let us consider next the problem of finding a maximum length path or circuit on G. This is equivalent to a minimization problem with distance matrix, $-D$, which has the Monge property. Hence, we know immediately that an optimal circuit has a pyramidal structure. This also turns out to be true for the optimal path. In addition, the pyramidal sequence is identified in both cases.

Lemma 1. *w_1 is not an end vertex in a maximum path.*

Proof: Simply flip any subsequence $w_1,..., w_s$ of the current solution, where w_s is an internal vertex. Suppose w_k was the adjacent vertex to w_s outside the subsequence. If w_1 is an end vertex, the difference between the new solution and the old one equals $(w_1 - w_s)w_k \geq 0$. ♦

Lemma 2. *A maximum path exists such that starting from w_1, both sides have non-increasing sequences of weights.*

Proof: Consider a given solution with subsequence $w_i, w_j, ... , w_r, w_s$ on one side of w_1, where $i < r < j$ and s. Flip the internal subsequence $w_j, ... , w_r$, so that w_r is now adjacent to w_i and w_j to w_s. The change in total length from the original solution is given by:

$$w_i w_r + w_j w_s - (w_i w_j + w_r w_s) = (w_i - w_s)(w_r - w_j) \geq 0.$$

A series of similar moves will yield a pyramidal structure without decreasing the length of the path. ♦

Property 4. *The maximum length traveling salesman path and circuit problems on G are both solved by the pyramidal sequence:*

$$w_N, w_{N-2}, w_{N-4},\ldots, w_1,\ldots, w_{N-5}, w_{N-3}, w_{N-1}. \tag{13}$$

Proof: First consider the path problem, and identify the first out-of-sequence weight starting from w_1 and expanding simultaneously on both sides of w_1. To illustrate, suppose we have the following case: ..., w_r, w_4, w_2, w_1, w_3, w_5, ..., where $r \neq 6$. There are two possibilities:

(i) w_6 is on the left side of w_1. We simply flip the subsequence $w_6,..., w_r$.

(ii) w_6 is on the right side of w_1. Now we exchange the left tail, ..., w_r, with the right tail, w_6,..., while flipping both.

In each case w_6 ends up in the required position; furthermore, it is easily shown that the total length of the path is not reduced. Using a series of similar moves, any sequence is converted to the pyramidal structure in (13) without reducing (and possibly increasing) the total path length. We conclude that (13) is a maximum length path on G.

The circuit problem is more straightforward. Since there is no left or right tail, we only use flips pertaining to case (*i*) to conclude that (13) also gives a maximum length circuit on G. ♦

4. APPLICATIONS

4.1. Military convoy

Consider a situation where *N* trucks will form a convoy to transport supplies from a base camp to an outlying military unit. The convoy will pass through a hostile territory where it may be subject to a missile attack. The job of the logistics officer in charge of the loading operation is to partition the supplies among the *N* trucks and then to arrange the sequence of the trucks in the convoy. Each subset S_t of supplies thus formed may be assigned a value q_t. The objective of minimizing the risk of loss from a single missile attack falls within the framework of the "buried treasure problem". If the S_t are unknown, the original problem (P1) is posed, and if they are given, the problem simplifies to (P2).

It is interesting to observe the built-in robustness of the sequence shown in (11) resulting from the alternate placement of high and low-valued subsets. Suppose, for example, that the convoy has 10 trucks with loads forming the sequence, $w_1 w_{10} w_3 w_8 w_5 w_6 w_7 w_4 w_9 w_2$, in accordance with (11). If the direction of travel is to the right, and the third and fourth trucks from the front (w_4 and w_7) are immobilized, the remaining convoy forms up directly as follows: $w_1 w_{10} w_3 w_8 w_5 w_6 w_9 w_2$.

Although this sequence is no longer optimal for the remaining loads, it is still a good solution since it is one flip $(w_5 w_6)$ away from optimality. This is true in general: if any two adjacent w_i are removed, the remaining sequence is at most one flip away from the optimal solution in (11). Unfortunately, this geometric property of the sequence is lost if only one w_i is removed.

4.2. Machine layout problem

Suppose there are *N* automated machine centres that are to be located at equally spaced distances along a linear aisle or a circular carousel in a flexible manufacturing system. The flow of material between any pair of centres is unknown or very difficult to predict accurately. However, we do have a good idea how busy each of these centres will be, so that the centres may be ordered and assigned indices w_i, where $w_i \geq w_{i+1}$, $i = 1, \ldots, N-1$. As a rough cut, the flow between centres i and j is assumed to be proportional to the product $w_i w_j$; thus the material handling cost for the pair (*i*, *j*) may be estimated as $w_i w_j d_{ij}$, where d_{ij} is the distance along the aisle or carousel between the two centres, which, of course, depends on their relative positions. The objective would be to locate the machine centres on the aisle or carousel in order to:

$$\text{minimize} \sum_{i<j}\sum w_i w_j d_{ij} .$$

It is readily shown that the solution given by (13) is at least locally optimal, since any pair-wise exchange of the w_i cannot improve this solution. This solution also appeals to the intuition by locating the busier centres close to each other.

4.3. Timing of new products

Suppose that a company has N new products ready to be launched in the market. Due to internal constraints and market conditions, the company policy requires that new products be launched one at a time at three-month intervals. The question then is one of timing.

Senior management has stipulated that the sequence of product launches should be as robust as possible to random shocks in the marketplace. More specifically, the planners are asked to design the schedule to factor in one severe economic downturn of five-month duration. Note that this scenario fits the model in (P2) with distance between caches now replaced by the time between successive product launches. The permutation given in (11), where the w_i are net present values, would provide a robust solution as requested by senior management.

5. A FINAL COMMENT

We return to the original problem (P1), and show that it is *NP-hard*, and thus, no efficient algorithm exists to solve it. Recall that this problem first requires the partitioning of m objects into N subsets before the assignment of subsets to caches which was investigated in a preceding section.

Suppose that $N = 2$ (only two caches). Then (P1) reduces to the following problem:

$$\min v_1^2 + v_2^2 + bv_1v_2$$

$$\text{s.t. } v_1 + v_2 = W,$$

where the total value W is given in (6) and $0 < b = (2l\text{-}a)/l < 1$. The minimum is obtained by $v_1 = v_2 = W/2$. It follows that, if we can solve (P1), we can also resolve the partition problem. But this latter problem is *NP-complete*. It may also be shown by a transformation from the 3-Partition problem that (P1) is NP-hard in the strong sense.

On the other hand, it is readily shown that the risk function B_1 is maximized by placing all the objects in any one of the N caches.

6. CONCLUSIONS

In this paper, we pose a hypothetical problem referred to as the “buried treasure problem”. This problem is shown to be NP-hard. However, an embedded sub-problem is identified as a special class of TSP for which there are known results. These results are extended to the traveling salesman path. Some possible applications of the model are also given.

Future research in this area includes: (i) solving the complete problem exactly or by heuristic methods, and (ii) generalizing the model to include interactions between non-adjacent pairs of locations.

REFERENCES

[1] Aizenshtat, V.S., and Kravchuk, D.N., "Minimum of a linear form on the set of all complete cycles of the symmetric group S_n", *Cybernetics*, 4 (1968) 52-53.

[2] Burkard, R.E., "Efficiently solvable special cases of hard combinatorial optimization problems", *Mathematical Programming*, 79 (1997) 55-69.

[3] Burkard, R.E., Cela, E., Rote, G., and Woeginger, G.J., “The quadratic assignment problem with an Anti-Monge and a Toeplitz matrix: easy and hard cases”, Technical report SFB-34, Institut für Mathematik B, Technische Universität Graz, 1995.

[4] Burkard, R.E., Deineko, V.G., Van Dal, R., Van Der Veen, J.A.A., and Woeginger, G.J., "Well-solvable special cases of the Traveling Salesman Problem: a survey", *SIAM Review*, 40 (1998) 496-546.

[5] Garey, M.R., and Johnson, D.S., *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman and Company, New York, 1979.

[6] Hallin, M., Melard, G., and Milhaud, X., "Permutational extreme values of autocorrelation coefficients and a Pitman test against serial dependence", *Ann. Statist*, 20 (1992) 523-534.

[7] Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G., and Shmoys, D.B., *The Travelling Salesman Problem: A Guided Tour of Combinatorial Optimization*, John Wiley & Sons, New York, 1990.

[8] Simeone, B., "An asymptotically exact polynomial algorithm for equipartition problems", *Discrete Applied Mathematics*, 14 (1986) 283-293.