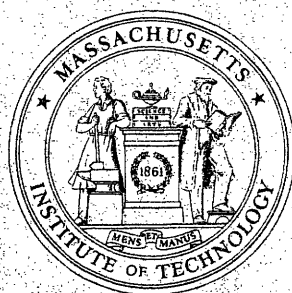


# OPERATIONS RESEARCH CENTER

working paper



**MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY**

A STATISTICAL APPROACH TO THE TSP

by

Bruce L. Golden

OR 052-76

April 1976

Supported in part by the U.S. Department of Transportation under contract DOT-TSC-1058, Transportation Advanced Research Program (TARP).

## ABSTRACT

This paper is an example of the growing interface between statistics and mathematical optimization. A very efficient heuristic algorithm for the combinatorially intractable TSP is presented, from which statistical estimates of the optimal tour length can be derived. Assumptions, along with computational experience and conclusions are discussed.

## Introduction

In this paper the "savings" technique for solving the vehicle routing problem, a generalization of the TSP, is used to solve large traveling salesman problems. Heuristic solutions are compared with expected values and statistical estimates of the optimal tour lengths. In particular, the Weibull distribution is used to model the distribution of heuristic solutions to the combinatorially explosive TSP. An algorithm is presented whose running time is especially encouraging.

The objectives of our work are three-fold and include efficiency, accuracy, and evaluation. That is, we hope to provide an extremely fast heuristic TSP algorithm which yields solutions which are within 5 or 10% of the optimal solution. In addition, based on the heuristic, we want to be able to estimate the optimal tour length in order to evaluate the heuristic more suitably.

The Clarke-Wright algorithm is a "greedy" heuristic algorithm which has been implemented to solve large-scale vehicle routing problems. Initially, each demand node is serviced separately from a specified central depot (node 1). Nodes  $i$  and  $j$  become linked according to the magnitude of the savings

$$s(i,j) \equiv d(1,i) + d(1,j) - d(i,j)$$

where  $d(i,j)$  is the distance from node  $i$  to node  $j$ . Figures I and II illustrate the savings formula. The basic motivation is as follows. If node  $i$  is an end-point of a tour and  $i$  can be linked feasibly with nodes  $j$  and  $k$ , and  $s(i,j) > s(i,k)$ , then linking  $i$  and  $j$  would be preferred (in the short range) to linking  $i$  and  $k$ , although possibly not in the optimal solution. The algorithm proceeds myopically, choosing the best feasible savings at each iteration from an

endpoint of the existing tour to a node not yet in the tour. See Clarke and Wright [3] and Golden, Magnanti, and Nguyen [8] for details. We have applied a modified version of the algorithm to a distribution system for an urban newspaper with an evening circulation exceeding 100,000. This problem contained nearly 600 drop points for newspaper bundles and was solved with 20 seconds of execution time on an IBM 370/168.

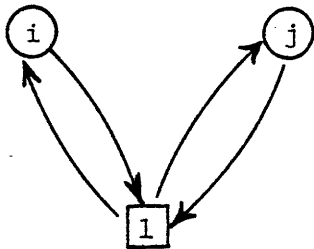


Figure I. Initial Setup.

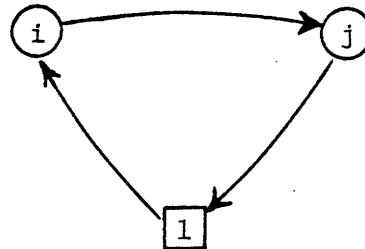


Figure II. Nodes i and j have been linked.

Golden, Magnanti, and Nguyen [8] have overcome some of the drawbacks of the original approach in their modified algorithm. The Clarke-Wright algorithm was designed initially to handle a matrix of real inputs, distances and savings. In dealing with a large problem, the number of required storage locations soon becomes unwieldy. Rather than consider all pairwise linkings, we can restrict our search to a small subset of the possible linkings, which we store in a list instead of a matrix. This is accomplished by superimposing an arbitrary grid over the nodes of the transportation network. Only the savings associated with arcs linking nodes in adjacent boxes are considered. Since the grid is arbitrary, care must be taken in the assignment of the parameters which determine the grid. At each step of the algorithm, we must determine the maximum savings. This comparison of savings is handled quickly and conveniently by partially ordering the data in a heap structure (see [8] for details) and updating the heap at each step after altering the routes.

We can use this savings approach for the solution of TSP's since a single vehicle of unlimited capacity leads to an appropriately defined VRP. In fact, the computer code can be streamlined to exploit this simplification, and extremely fast running times result. However, in some cases we sacrifice accuracy for speed to an unacceptable degree. Since we are building a hamiltonian circuit, we can consider any node to be the origin (this cannot be done in solving the VRP). If we consider several independent origins we can produce several independent heuristic solutions and simply choose the minimum length tour as our heuristic TSP solution. We gain accuracy, lose little in speed, and hopefully obtain a means for evaluating our heuristic solution. The evaluation rests on the assumption that the independent heuristic solutions can be modeled as generated from a Weibull distribution, an issue we will take up later. The assumption of independent heuristic solutions needs, perhaps, some justification. We can generate  $N$  origins randomly from a network of  $n$  nodes with or without replacement. When we choose with replacement, the heuristic solutions will be independent. However, we prefer to select  $N$  distinct origins. When we choose without replacement we obtain more information since each new origin generates a new heuristic solution, and if the ratio  $N/n$  is small, then the heuristic solutions will be more or less independent. That is, the difference between sampling with replacement and without replacement is significant only when the population we are sampling from contains relatively few members. When the ratio  $N/n$  is small the selection and nonreplacement of a particular node has a negligible effect on the probabilities of selecting additional nodes.

An underlying issue involves the evaluation of heuristic solutions to NP-complete problems, such as the TSP. Rosenkrantz, Stearns, and Lewis [20]

have obtained, through elegant combinatorial arguments, some interesting (although not especially reassuring) worst-case performance bounds for approximate algorithms for the TSP when the triangle inequality holds. For the nearest neighbor method, the worst-case ratio of the length of the obtained tour to the length of the optimal tour increases logarithmically with the number of nodes. The nearest insertion method and the cheapest insertion method have worst-case ratios which approach 2 as the number of nodes increases. We have shown that the worst-case ratio of the Clarke-Wright solution (from a single origin) to the optimal solution can be arbitrarily large as the number of nodes increases [7].

#### Statistical Background

The Weibull family of distributions has been studied extensively in recent years, as evidenced by the journal Technometrics. It is especially useful in problems of life testing and reliability where the life length may be bounded from below (See Barlow and Proschan [1]). This family is, in some sense, a more generalized form of the exponential distribution since it has three parameters and can be reduced to the exponential distribution with the proper choice ( $c=1$ ) of one of them. Gumbel [9] refers to the Weibull family as the Type III asymptotic distribution of extreme values. Fisher and Tippett [6], in their fundamental 1928 paper, were the first to derive the three asymptotic distributions. We will come back to this later. The probability density function for the Weibull distribution is

$$f_x(x_0) = \left(\frac{c}{b}\right) \left(\frac{x_0 - a}{b}\right)^{c-1} \exp \left[ -\left(\frac{x_0 - a}{b}\right)^c \right] \text{ for } x_0 \geq a \geq 0, b > 0, c > 0,$$

where  $a$  is the location parameter,  $b$  is the scale parameter, and  $c$  is the shape parameter. It is often easier to work with the cumulative Weibull distribution

given by

$$\text{Prob } \{x \leq x_0\} = 1 - \exp \left[ - \left( \frac{x_0 - a}{b} \right)^c \right].$$

Note that the random variable  $(x-a)^c$  has an exponential distribution with expected value  $b^c$  since

$$\text{Prob } \{(x-a)^c \leq x_0\} = 1 - \exp \left[ \frac{-x_0}{b^c} \right].$$

McRoberts [18], in dealing with combinatorially explosive plant-layout problems, introduced the idea of matching the distribution of heuristic solutions with a Weibull distribution. He further suggests that other combinatorial problems might be approached in a similar manner. Unfortunately, his arguments were substantially weakened by the fact that he treats his intermediate heuristic solutions as a set of independent observations from a parent distribution despite the fact that there exists a clear interdependence among these solutions. In part, this paper is an outgrowth of his work.

Consider  $N$  samples, each of size  $m$ , taken from a parent population which is bounded from below. In each sample there is a smallest value  $x_i$  and the smallest value  $v$  in  $Nm$  observations is the smallest of the  $N$  smallest values  $x_i$ , i.e.,  $v = \min\{x_i | 1 \leq i \leq N\}$ . All sampling is independent. Fisher and Tippett [6] demonstrated that when  $m$  is very large the distribution of  $x_i$  approaches what we now call the Weibull distribution. Gumbel[9] points out that in studying one extreme, no assumption need be made regarding the behavior of the initial distribution at the other extreme. The Weibull distribution is independent of the parent distribution, so we do not need to know the distribution from which the samples are generated. Weibull was an engineer who later derived, in an empirical way, the same distribution and applied it



to the analysis of dynamic breaking strength [22].

Intuitively, it seems reasonable that the distribution of heuristic solutions could be Weibull. Suppose there are  $n$  nodes. Then, the parent population consists of  $(n-1)!/2$  tours, with lengths bounded from below by the length of the optimal tour. Each independent heuristic solution implicitly is a local minimum from a large number  $m$  of possible tours. With this in mind, we can perform the proposed algorithm from various origins to obtain heuristic solutions (tour lengths)  $x_1, x_2, \dots, x_N$ . Next, using maximum likelihood estimation we might find best estimates for  $a, b$ , and  $c$ .

The likelihood function is given by

$$\begin{aligned} L(x_1, x_2, \dots, x_N; a, b, c) &= \left(\frac{c}{b}\right) \left(\frac{x_1 - a}{b}\right)^{c-1} \exp \left[ - \left(\frac{x_1 - a}{b}\right)^c \right] \\ &\quad \dots \left(\frac{c}{b}\right) \left(\frac{x_N - a}{b}\right)^{c-1} \exp \left[ - \left(\frac{x_N - a}{b}\right)^c \right] \\ &= \left(\frac{c}{b}\right)^N \left(\frac{1}{b}\right)^{Nc-N} \left\{ (x_1 - a) \dots (x_N - a) \right\}^{c-1} \exp \left[ - \sum_{i=1}^N \left(\frac{x_i - a}{b}\right)^c \right]. \end{aligned}$$

Since  $L(\theta)$  and  $\ln(L(\theta))$  have their maximum at the same value of  $\theta$ , we maximize the natural logarithm of the likelihood

$$\ln L = N \ln c - Nc \ln b + (c-1) \sum_{i=1}^N \ln(x_i - a) - \sum_{i=1}^N \left(\frac{x_i - a}{b}\right)^c.$$

If we now equate  $\frac{\partial \ln L}{\partial a} = \frac{\partial \ln L}{\partial b} = \frac{\partial \ln L}{\partial c} = 0$ , we obtain the following maximum likelihood equations:

$$(1) \quad \frac{\partial \ln L}{\partial a} = -(c-1) \sum_{i=1}^N \left(\frac{1}{x_i - a}\right) + \left(\frac{c}{b}\right) \sum_{i=1}^N (x_i - a)^{c-1} = 0;$$

$$(2) \quad \frac{\partial \ln L}{\partial b} = \frac{-NC}{b} + \frac{c}{b^{c+1}} \sum_{i=1}^N (x_i - a)^c = 0;$$

$$(3) \quad \frac{\partial \ln L}{\partial c} = \frac{N}{c} - N \ln b + \sum_{i=1}^N \ln (x_i - a) - \sum_{i=1}^N \left( \frac{x_i - a}{b} \right)^c \ln \left( \frac{x_i - a}{b} \right) = 0.$$

Equation (2) yields 
$$b = \left\{ \frac{\sum_{i=1}^N (x_i - a)^c}{N} \right\}^{\frac{1}{c}} \quad (*)$$

which upon substitution into equations (1) and (3) (see Markham [16]) gives

$$(4) \quad -(c-1) \sum_{i=1}^N \left( \frac{1}{x_i - a} \right) + NC \sum_{i=1}^N (x_i - a)^{c-1} / \sum_{i=1}^N (x_i - a)^c = 0 \quad \text{and}$$

$$(5) \quad \frac{N}{c} + \sum_{i=1}^N \ln (x_i - a) - N \sum_{i=1}^N (x_i - a)^c \ln(x_i - a) / \sum_{i=1}^N (x_i - a)^c = 0.$$

One approach is to solve the two markedly nonlinear constraints (4) and (5) by numerical methods. With estimates for a and c, then compute b directly from (\*).

A second approach involves maximizing ln L directly, subject to the constraints  $b > 0$ ,  $c > 0$ . Barrier methods are often applicable to problems of this form. See Fiacco and McCormick [5] for a thorough treatment of penalty and barrier function techniques.

We have chosen a less cumbersome third line of attack based on the least squares concept. The Weibull cumulative distribution function provides the initial equation

$$\exp \left[ - \left( \frac{x_0 - a}{b} \right)^c \right] = 1 - F(x_0).$$

Upon taking logarithms twice, we obtain another equation,

$$c \ln (x_0 - a) - c \ln b = \ln (-\ln(1 - F(x_0))) \quad (6)$$

which is the equation of a straight line with dependent variable  $\ln(-\ln(1 - F(x_0)))$  and independent variable  $\ln(x_0 - a)$ . If we fix  $a$ , and use least squares analysis to estimate  $\alpha$  and  $\beta$  in

$$\ln(-\ln(1 - F(x_0))) = \alpha \ln(x_0 - a) + \beta,$$

then  $b$  and  $c$  can also be estimated. Moreover, we can perform a line search on  $a$  to determine the set of parameter values which yields the largest correlation coefficient. A range of possible values for  $a$  is obtained easily since we know that  $a$  must be less than  $v = \min\{x_i \mid 1 \leq i \leq N\}$ .  $F(x_0)$  is estimated from the sample data. Since the correlation coefficient measures the strength of the linear relationship between the dependent and independent variables, if the best correlation coefficient is not close to unity, in absolute value, then we should begin to question the null hypothesis that the heuristic solutions are Weibull distributed.

This least squares approach has been implemented and seems to perform remarkably well. From experiments, the parameter estimates that it produces solve the maximum likelihood equations (1)-(3) almost exactly. In addition, the correlation coefficient is always very close to 1.

With parameters estimated, the well-known Kolmogorov-Smirnov test can be used to test the null hypothesis. This test compares the theoretical and sample cumulative distribution functions. If the largest absolute vertical deviation is beyond a critical value that depends on the significance level,

then the hypothesis that the two distributions are identical must be rejected. If, indeed, the null hypothesis cannot be rejected, we have an estimate for  $a$ , the optimal TSP solution. It should be noted that since the parameters  $a, b$ , and  $c$  are estimated from the data, the Kolmogorov-Smirnov test is an approximate test. The Kolmogorov-Smirnov statistic has traditionally been applied in situations where the theoretical cumulative distribution function is completely specified, i.e., all parameters are known. In many practical situations, however, some or all of the parameters are unknown in which case they must be estimated from the sample data; an approximate test results. Recent research in mathematical statistics has begun to focus on the exact nature of this approximation (see Stephens [21]).

#### Expected Length of an Optimal Tour

Beardwood, Halton, and Hammersley [2] derive the asymptotic expected length of an optimal traveling salesman tour for a special class of networks. In addition, they prove that the variance of the length goes to zero as  $n$  becomes large. For an  $n$  node problem ( $n$  large) where the nodes are distributed randomly and uniformly over some arbitrary area of  $S$  units, the expected length of the optimal TSP tour,  $L(n,S)$ , is given by

$$L(n,S) = K\sqrt{n} \sqrt{S} . \tag{7}$$

In their excellent book, Eilon, Watson-Gandy, and Christofides [4] perform simulations which indicate that  $K = .75$  approximately; the expected length formula is reasonably accurate with  $K = .75$  although one could argue that it consistently underestimates the optimal tour length (see Fig. 8.18 [4]). This formula is important since for large  $n$  we cannot solve TSP's exactly.

We now have two means for estimating the optimal TSP solution. Our goal is to test the Weibull estimate against the expected value formula (7) for moderate-sized problems (70 to 130 nodes) where the nodes are distributed randomly over an area in euclidean space. Of course, it can be argued that for very large networks when these conditions are not satisfied the expected length formula still provides an excellent approximation. This, however, is difficult to substantiate. Moreover, the statistical estimate is problem-dependent, and so it takes into account, rather than ignores, the possible nonrandomness of nodes.

#### Discussion of Algorithm

Every arc that we consider for linking in our modified Clarke-Wright algorithm has a corresponding savings relative to a particular origin. We order these savings from greatest to least on a heap and starting from the top of the list we link nodes  $i$  and  $j$  where  $s_{ij}$  represents the current maximum feasible savings. We continue until a tour on  $n$  nodes is formed. Linking nodes  $i$  and  $j$  is feasible so long as neither  $i$  nor  $j$  are interior to a sub-tour.

Now, when a number of distinct origins are generated for each problem, we can eliminate redundant calculations in the following way. First, define the arbitrary grid judiciously (set  $DIV$ ). Given the  $x$  and  $y$  coordinates of the nodes in the network, the grid is divided into  $DIV^2$  equally-sized rectangles so that each node  $i$  has box coordinates  $BX(i)$  and  $BY(i)$ . Arcs with nodes in adjacent boxes are considered for linking; these arcs are stored along with their distances. In other words, if  $|BX(i) - BX(j)| > 1$  or  $|BY(i) - BY(j)| > 1$  then arc  $(i,j)$  is ignored. From computational experience, which will be discussed later in this paper, the number of nodes in the network suggests an

appropriate range of values for the parameter DIV. As the origins are varied, only the savings values change. The computer code reflects this observation. For each different origin  $o$ , we must redefine our savings function accordingly

$$s_{ij}(o) \equiv d(o,i) + d(o,j) - d(i,j) \quad (8)$$

and use the savings to construct a tour as discussed previously.

### Computational Results

We have performed extensive computational tests essentially to address the following three questions:

- (i) What is the efficiency and accuracy of the proposed heuristic solution technique?
- (ii) Can the Weibull hypothesis be justified statistically?
- (iii) How do the estimated solutions and the expected solutions compare?

The results obtained tend to confirm our intuition; Tables I and II display the numerical findings.

Table I focuses on questions (i) and (ii). Networks of from 70 to 130 nodes were generated randomly in a square of area 10,000. In Table I, for each value of  $n$  ( $n = 70, 80, \dots, 130$ ) a single network was generated. The savings heuristic was applied from  $N$  distinct origins with a grid of  $DIV^2$  boxes to obtain a final heuristic solution. Running times are remarkably fast, ranging from 16 to 40 seconds in total execution time on an IBM 370/168. These run times include all input and output operations (time spent generating networks is also included). The final heuristic solutions are within 4-10% of the expected solution, except in one instance. The parameters  $a, b$ , and  $c$  have been estimated and the location parameter  $a$  seems to be slightly above the

expected solution in general. Finally, in all cases, the observed Kolmogorov-Smirnov statistics fall far below the critical value at the .05 level of significance.

Table II deals with the third question. Here, for each value of  $n$ , five networks have been generated as before, in order to study average behavior. The average estimated solution is compared with the expected solution. In all cases the average estimated solution is above and within about five percent of the expected solution. This represents a fairly close fit. Furthermore, if  $K$  is increased to .76, as perhaps it should be (although .75 is certainly more convenient), the fit becomes still better.

For all problems in Table I, estimating the location parameter was especially easy. Plotting correlation coefficient as a function of location parameter we find a unimodal function as indicated in Table III where the entries correspond to the case  $n = 130$  from Table I.

Next, we decided to construct a test network where the expected length formula was clearly not applicable. In a square area of 8 units, we let the 81 node locations in our network be at the integer lattice points  $(i,j)$  contained in the region. The proposed heuristic was applied and the observed Kolmogorov-Smirnov statistic was found to fall far below the critical level, as in the Table I experiments. In addition, we solved the 25 - city problem posed by Held and Karp [10] who conjectured an optimal solution of length 1711. units. Little et al, later verified this conjecture [15]. The expected length formula does not hold in this case either. In a second of computer time, tours from ten distinct origins were generated. Our algorithm's solution of 1750. units is about 2.2% above the optimal solution. The estimated optimal solution under the Weibull assumption was found to be 1725. units - less than

Table I. Computational Results. Each entry refers to one network.

n	N	DIV	Running Time (seconds)	Expected Solution (1)	Heuristic Solution (2)	Approximate Deviation $\frac{(2)-(1)}{(1)}$	Parameter Estimation			Observed K-S Statistic	Critical Value
							a	b	c		
70	25	4	16	627.75	659.39	5.0%	650.	37.78	1.95	.083	.27
80	25	4	21	670.50	700.46	4.5%	695.	37.77	1.56	.097	.27
90	25	5	19	711.75	747.24	5.0%	740.	35.74	2.01	.111	.27
100	25	5	24	750.00	785.73	4.8%	720.	117.42	4.83	.114	.27
110	30	6	33	786.75	832.44	5.8%	770.	111.65	6.18	.106	.24
120	30	6	40	821.25	915.36	11.4%	870.	82.30	4.90	.086	.24
130	30	7	37	855.00	917.56	7.3%	900.	91.72	2.00	.084	.24



Table II. Computational Results. Each entry is an average over five networks.

n	N	DIV	Expected Solution (1)	Average Estimated Solution (2)	Average Heuristic Solution (3)	Approx. Deviation $\frac{(2)-(1)}{(1)}$	Approx. Deviation $\frac{(3)-(1)}{(1)}$
70	25	4	627.75	643.	690.	2.4%	9.9%
80	25	4	670.50	695.	727.	3.6%	8.4%
90	25	5	711.75	740.	763.	3.9%	7.2%
100	25	5	750.00	774.	804.	3.2%	7.2%
110	30	6	786.75	805.	836.	2.3%	6.2%
120	30	6	821.25	837.	879.	1.9%	7.0%
130	30	7	855.00	899.	917.	5.1%	7.2%

Table III. Correlation coefficient as a function of location parameter.

Location Parameter	Correlation Coefficient
850	.970
855	.971
860	.972
865	.973
870	.974
875	.975
880	.976
885	.977
890	.979
895	.980
900	.982
905	.981
910	.977
915	.949

1% away from optimality. The observed Kolmogorov-Smirnov statistic was .071 against a critical value of .410 at the .05 level of significance. These experiments have signaled that, indeed, the Weibull assumption can be exploited in more general situations than the expected value result.

Finally, we sought to determine the discriminating power of the Kolmogorov-Smirnov test. The normal distribution provides a fairly accurate approximation to the Weibull in this problem, especially in terms of central tendency. We experimented with problem #24 from Krolak et al. [13] (some of our results will be mentioned in the next section) first assuming an underlying Weibull distribution, then an underlying normal distribution. The observed Kolmogorov-Smirnov statistic is .056 under the Weibull hypothesis, and .130 under the normal hypothesis - quite a difference. Although neither assumption would be rejected at the .05 level of significance, the Weibull assumption yields a much closer fit.

#### Other TSP Heuristics

The interchange algorithm of Lin and Kernighan [14] is probably the most effective procedure available for generating optimum and near-optimum solutions to the symmetric TSP. The heart of their procedure involves a transformation whereby  $k$  edges in the current tour are replaced by  $k$  other edges, yielding a better tour. They claim to have solved 100 node problems exactly with 99 per cent confidence in 3-4 minutes running time on a GE 635. In contrast, our procedure solves 130 node problems to within approximately 7% of the optimal solution in about 37 seconds.

Since Lin and Kernighan and Krolak et al. have each obtained the solution 21282. for problem #24 [13], it is believed to be the optimal solution. In applying our TSP heuristic, we obtained a solution of 21978. in 22 seconds

of IBM 370/168 execution time which is only 3.3% away from optimality. In addition, our estimate for the location parameter was 21650., about 1.7% above the optimal solution.

Our heuristic is faster than the Lin-Kernighan heuristic although not as accurate, and much faster than the Krolak et al. man-machine approach. For problems involving more than 110 nodes the computer storage requirements become too excessive for the present version of the Lin-Kernighan heuristic. Modifications are indicated (although not implemented) in their paper. Storage is not a problem with the proposed algorithm since we are selective in the arcs which we choose to consider for linking, and larger problems can be handled without difficulty. It is also conceivable that a hybrid approach in which a solution from our heuristic becomes an initial feasible solution for the interchange algorithm might be very successful.

### Extensions

In the event that distances are not euclidean, we can store the  $p$  nearest neighbors to each of the  $n$  nodes in a matrix. A shortest path algorithm can be used to find these nearest neighbors. Next, the data can be arranged in an  $n$  by  $3p$  matrix  $T$  where for  $L = 1, 2, \dots, p$ :

$T(I, 3L-2)$  = the node adjacent to node  $I$ ,

$T(I, 3L-1)$  = the distance between nodes  $I$  and  $T(I, 3L-2)$ ,

$T(I, 3L)$  = the savings obtained by linking  $I$  and  $T(I, 3L-2)$ .

This storage scheme is similar to the one developed by Williams [23] for shortest paths. For each node  $I$ , we can order the corresponding savings from largest to smallest via heap structures and proceed as before.

## Conclusions

If we examine the five 100 node problems in Krolak et al. [13] (we presume all nodes have been generated randomly in a rectangle of 2000 by 4000 units for each problem) and if we conjecture that the Lin-Kernighan solutions are optimal, then we can compare the average optimal solution 21508. with the expected optimal tour length. Setting  $K$  to .76 rather than .75 we obtain an expected length of about 21500. The limited evidence suggests that .75 is probably too conservative (low) for  $K$ . On the other hand, there may be a complex underlying bias mechanism which causes our location parameter estimates to consistently overestimate by 1 or 2%. Notice from Table II that there are no trends towards greater disparity as  $n$  increases.

Recent results in complexity theory indicate that many network optimization problems such as the TSP are inherently difficult to solve. In fact, it seems unlikely that polynomial algorithms can be obtained for exact solution to these problems. With this in mind, heuristic algorithms have become increasingly important. In this paper, we have presented an efficient and accurate heuristic algorithm for approximate solution of the TSP. Of course, if an improved solution is desired, we can determine some or all of the  $n-N$  remaining heuristic solutions in a reasonable amount of computer time. In addition, we have provided a statistical approach for estimating the optimal solution, which will help in assessing deviations from optimality. On the basis of our computational results we cannot reject the null hypothesis that, indeed, an underlying Weibull distribution is at work. On the other hand, as with all statistical arguments, we cannot arrive at an absolutely firm conclusion. The suggested statistical approach, along with sharp lower bounds from lagrangean relaxation [11],[12], and combinatorial worst-case ratio analysis

[20], is yet another means for evaluating heuristics for hard combinatorial optimization problems.

Acknowledgments

The author would like to thank Professors T. Magnanti, J. Ferreira, and R. Welsch, and Mr. R. Wong for their valuable suggestions.

References

- [1] R. Barlow and F. Proschan, Mathematical Theory of Reliability, John Wiley & Sons, New York (1965).
- [2] J. Beardwood, J. Halton, and J. Hammersley, "The Shortest Path through Many Points," Proc. Camb. Phil. Soc., 55, 299 (1959).
- [3] G. Clarke and J. Wright, "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points," Oper. Res., 12, 568 (1964).
- [4] S. Eilon, C. Watson-Gandy, and N. Christofides, Distribution Management, Griffin, London (1971).
- [5] A. Fiacco and G. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley, New York (1968).
- [6] R. Fisher and L. Tippett, "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample," Proc. Camb. Phil. Soc., 24, 180(1928).
- [7] B. Golden, "Recent Developments in Vehicle Routing," in preparation.
- [8] B. Golden, T. Magnanti, and H. Nguyen, "Implementing Vehicle Routing Algorithms," M.I.T. Operations Research Center Technical Report, No. 115 (September 1975).
- [9] E. Gumbel, Statistics of Extremes, Columbia University Press, New York (1958).
- [10] M. Held and R. Karp, "A Dynamic-Programming Approach to Sequencing Problems," J. SIAM, 10, 196-210 (1962).
- [11] M. Held and R. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees," Oper. Res., 18 (6), 1138-1162 (1970).
- [12] M. Held and R. Karp, "The Traveling Salesman Problem and Minimum Spanning Trees: Part II," Math. Prog., 1, 6-25(1971).
- [13] P. Krolak, W. Felts, and G. Marble, "A Man-Machine Approach Toward Solving the Traveling Salesman Problem," CACM, 14, 327-334(1971).
- [14] S. Lin and B. Kernighan, "An Effective Heuristic Algorithm for the Traveling Salesman Problem," Oper. Res., 21, 498-516 (1973).
- [15] J. Little, K. Murty, D. Sweeney, and C. Karel, "An Algorithm for the Traveling Salesman Problem," Oper. Res., 11 (5), 972-989(1963).
- [16] R. Markham, "Some Aspects of the Weibull Distribution," Masters Thesis, Department of Mathematical Statistics, University of South Africa (1974).

- [17] K. McRoberts, "Optimization of Facility Layout," PhD Thesis, Iowa State University of Science and Technology, Ames, Iowa (1966).
- [18] K. McRoberts, "A Search Model for Evaluating Combinatorially Explosive Problems," Oper. Res., 19, 1331 (1971).
- [19] J. Ravis, "Life Testing: Estimating the Parameters of the Weibull Distribution," IEEE International Convention Record (1963), 18-33.
- [20] D. Rosenkrantz, R. Stearns, and P. Lewis, "Approximate Algorithms for the Traveling Salesperson Problem," Proc. of the 15th IEEE Symp. on Switching and Automata Theory, 33-42 (1974).
- [21] M. Stephens, "Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters," Ann. Statist., 4(2), 357-369 (1976).
- [22] W. Weibull, "A Statistical Distribution Function of Wide Applicability," J. Appl. Mech., 18, 293 (1951).
- [23] T. Williams, "The Shortest Path Problem: An Intermediate Stage Results Approach," presented at November 1975 ORSA Meeting, Las Vegas, Nevada.
- [24] R. Winkler and W. Hays, Statistics: Probability, Inference, and Decision, Holt, Rinehart, and Winston, New York (1975).