Algorithms for the Universal and *A Priori* TSP

Frans Schalekamp [*]
Cornell University
Ithaca, NY

David B. Shmoys
Cornell University
Ithaca, NY

Abstract

We present two simple results for generalizations of the traveling salesman problem (TSP): For the universal TSP, we show that one can compute a tour that is universally optimal whenever the input is a tree metric. A (randomized) $O(\log n)$-approximation algorithm for the *a priori* TSP follows as a corollary.

Keywords: Traveling Salesman Problem, A Priori Optimization, Universal Optimization, Metric Embedding

[*]Corresponding author. Email address: fms9@cornell.edu. Postal address: School of ORIE, Cornell University, 288 Rhodes Hall, Ithaca, NY 14853.

1 Introduction

Consider the following problem: A delivery person needs to serve clients each day. Her (potential) client set is fixed and known, but each day only a subset of the potential client set needs to be served. Instead of reoptimizing every day, the delivery person would like to have a tour of the potential client set, such that if she travels the subtour induced from the complete tour for that day's subset of clients (i.e., she visits the clients in the same order as the complete tour, shortcutting the tour where clients are absent), then she does not travel much more than had she used the optimal tour for this particular subset of clients.

More precisely, two different problems can be considered. One is an *adversial* model, where any of the subsets of clients is possible, and the objective is to minimize the maximum ratio, over all possible client subsets, of the length of the induced subtour to the length of the optimal tour for that subset. (Another way to view this is that the optimizer announces the complete tour, and next the adversary can decide which subset of clients actually need service). This is known as the *universal traveling salesman problem.*

A second model is a *probabilistic* model. Here one assumes that there is some probability distribution on subsets of clients and the objective is to minimize the expected length of the induced subtour. This problem is known as the *a priori traveling salesman problem.* (Note that there are complexity issues arising from specifying the probability distribution.)

In this paper we will show that when the client distances form a *tree metric*, we can, quite surprisingly, solve both problems optimally, and we can even solve the probabilistic model without any knowledge of the probability distribution.

This special case of tree metrics is particularly significant, because any metric can be (probabilistically) embedded in tree metrics with expected distortion $O(\log n)$ (where n is the number of points in the metric), as shown by Fakcharoenphol, Rao and Talwar [10] (building on Bartal's work [1]). As a corollary we therefore get that the *a priori* TSP on *any* metric space can be solved within $O(\log n)$ of optimal, even without any knowledge of the distribution on the subsets. It is interesting to note that this is the same guarantee as Platzman and Bartholdi's result for the universal TSP problem in the Euclidean plane [2, 17].

The universal TSP was motivated by getting good solutions for the "Meals On Wheels" program of Senior Citizen Services Inc., "which delivers prepared lunches to people who are unable to shop or cook for themselves"[3]. The list of clients that need service in this particular application is quite volatile "...because of the nature of the clients: most are elderly or ill. They may die, or recover from illness, or receive care elsewhere...”[3]. The approximation algorithm of [2, 17] uses a spacefilling curve to map points from the plane to the unit line (interpreted as a circle), on which a traveling salesman tour is trivial to solve. They proved a performance guarantee of $O(\log n)$, and conjectured that it was really $O(1)$. Bertsimas and Grigni [5] disproved this conjecture for Bartholdi and Platzman's algorithm, exhibiting a (family of) counterexample(s) for which $O(\log n)$ is tight up to a multiplicative constant.

Recently the universal TSP has again become an object of study, with [15] giving an algorithm for which the induced tour for any subset is within an $O(\log^4 n/\log\log n)$ factor of optimal, and even more recently [11] improves this bound to $O(\log^2 n)$, using the ideas of [10]. A lower bound for the 2-dimensional Euclidean universal TSP (i.e., independent of an algorithm) of $\Omega(\sqrt[6]{\log n/\log\log n})$ was presented in [12].

Jaillet introduced the notion of *a priori* optimization, in particular the *a priori* TSP [13, 14]. Except for asymptotic results on points distributed independently and uniformly on the plane [14, 6, 4], not much is known. (Note, of course, that any guarantee for the universal TSP implies the same guarantee for the *a priori* TSP.)

In [8] and [7] a related question was studied, namely, for what metrics does there exist a universal TSP solution such that the induced subtours are actually optimal solutions to the smaller problems? Or phrased differently, what metrics give a universal TSP solution with optimal ratio equal to 1? This property is called the master tour property and it turns out that the class of metrics that possess this property is fully described and known as *Kalmanson metrics*, as shown in [8].

2 A Priori and Universal TSP

We start by introducing some notation needed to present our results. For the traveling salesman problem, the input consists of a finite set of points $V = \{1, 2, \ldots, n\}$ and a “distance” function $d : V \times V \rightarrow Q^{\geq 0}$; we wish to find a tour τ, given by a permutation of all points in V, such that the length of the tour, $d(\tau) = \sum_{i=1}^{n-1} d(\tau_i, \tau_{i+1}) + d(\tau_n, \tau_1)$ is minimized.

For the *a priori* and universal TSP, only a subset $S \subseteq V$ will be “active”. For a tour τ of V, we define $\tau(S)$ to be the tour restricted to S, i.e. $(\tau(S))_i = \tau_j$ where j is such that $\tau_j \in S$ and $\#\{\tau_k \in S, k \leq j\} = i$ (where $\#\{N\}$ denotes the cardinality of the set N). For each $S \subseteq V$, we also let $OPT(S)$ denote the optimal tour on S and let $d(OPT(S))$ denote its length.

The input and output for the *universal TSP* are the same as for the TSP: we are given V and d, and we want a tour τ. The objective, however, now is to minimize, over all subsets $S \subseteq V$, the ratio of the length of the tour induced by the permutation τ on S, divided by the length of the optimal tour on S, i.e. to minimize $\max_{S \subseteq V}[d(\tau(S))/d(OPT(S))]$.

For the *a priori TSP*, the input is again a set of points V, and a distance function d, but there is also a probability distribution specified over the subsets of V. (For our results, it turns out that we do not need any knowledge of this distribution.) The goal now is to find a tour τ that minimizes the *expected length* (with respect to the probability distribution on subsets of V) of the tour induced by τ; i.e., to minimize $E_S[d(\tau|S)]$.

Let V be a finite set of points. A distance function $d : V \times V \rightarrow Q^{\geq 0}$ is called a *metric* if (1) $d(i,i) = 0$ for all $i \in V$ and (2) $d(i,j) \leq d(i,k) + d(k,j)$ for all $i,j,k \in V$ (the triangle inequality). (To be precise, this is the definition of a semimetric. For a semimetric to be a metric we also want $d(i,j) > 0$ if $i \neq j$, but this is of no importance for this paper.) Furthermore, we say that d is a *tree metric* if there exists a tree $T = (V_T, E_T)$ with nodes $V_T \supseteq V$ and lengths associated with the edges, such that $d(i,j)$ is exactly equal to the length of the unique path in T from i to j, for all $i,j \in V$. For each subset $S \subseteq V$, one can obtain an *induced tree* $T(S)$, which is the union over all pairs of nodes $i,j \in S$ of the edges (and nodes) in the path between i and j. We can now state and prove our results.

Lemma 1 *Consider an input to the TSP given by a tree metric on V that is realized by the tree T; then for any subset $S \subseteq V$, the length of any tour on S is at least twice the total length of the edges in the induced tree $T(S)$.*

Proof. Take any traveling salesman tour τ on S. We can transform this to a walk using the edges of $T(S)$, by inserting, between each pair of consecutive points i and j in τ, the unique path in T connecting i and j; this does not change the length of the tour. So we can view any tour as a walk that only uses the edges of the tree $T(S)$ (multiple times). However, consider any edge in $T(S)$; if we delete this edge, then we separate the tree $T(S)$ into two components, thereby partitioning S into two non-empty subsets S_1 and S_2. Thus, the walk must contain this edge at least twice in the tour, at least once "entering" S_1 and at least once "leaving" it. ■

For any tree T, we can consider the walk formed by traversing "around the outside" of T: we shall call this walk $\mathsf{trav}(T)$. Such a walk can be shortcut in a number of ways to obtain a tour; for a tree metric, all such tours are the same length, and by Lemma 1 are all optimal solutions to the TSP. It is easy to see that if we start with one such tour τ, and consider the tour $\tau(S)$ for any subset $S \subseteq V$, then this tour can be obtained by shortcutting the walk $\mathsf{trav}(T(S))$. Hence $\tau(S)$ is an optimal solution to the TSP input induced by S, and we have shown the following result.

Theorem 2 *For any tree metric, any shortcutting of the walk* $\mathsf{trav}(T)$ *yields an optimal solution to the universal TSP, and therefore an optimal solution to the* a priori *TSP.*

Note that the results above are also implied by the fact that tree metrics are a subclass of Kalmanson metrics, and the fact that Kalmanson metrics have the master tour property as proved in [8]. The proofs here are much simpler, as we just consider tree metrics.

Corollary 3 *There is a polynomial-time randomized algorithm that, for any metric input* d *to the TSP on* n *points, computes a tour* τ *such that for each subset* $S \subseteq V$*, the expected cost of the tour* $\tau(S)$ *is* $O(\log n)d(OPT(S))$ *(where the expectation is with respect to the random choices of the algorithm).*

Proof. Use (Bartal's) stochastic tree embedding [1, 9, 10]: this gives a distribution over tree metrics, d^T, such that $d^T(i,j) \geq d(i,j)$ for all $i,j \in V$, and $E_T[d^T(i,j)] = O(\log(n))d(i,j)$ for all $i,j \in V$. (We will use E_T to denote the expectation with respect to this distribution over the tree metrics.) For each subset $S \subseteq V$ and each T in the distribution, note that $d^T(\mathsf{trav}(T(S)) \leq d^T(OPT(S))$, since $OPT(S)$ is the optimal tour for S with respect to the original distance metric d, whereas the traversal yields the optimal tour for d^T. Therefore,

$$E_T[d^T(\mathsf{trav}(T(S)))] \leq E_T[d^T(OPT(S))]$$

for each $S \subseteq V$. We can now conclude that

$$E_T[d(\mathsf{trav}(T(S)))] \leq E_T[d^T(\mathsf{trav}(T(S)))] \leq E_T[d^T(OPT(S))] = O(\log n)d(OPT(S)),$$

where the first and third inequalities are due the properties of the embedding (non-shrinking, and in expectation expanding by at most $O(\log n)$). ■

Corollary 4 *There is a randomized $O(\log n)$-approximation algorithm for the a priori TSP on any metric space with n points — even without specifying the probability distribution on the subsets.*

Proof. The result follows rather directly from Corollary 3. Denote by p_S the probability that subset S "occurs". The expected cost (with respect to the random tree metric) of our *a priori* TSP solution (which involves an expectation with respect to the choice of S) is

$$\begin{aligned} E_T[\sum_S p_S d(\mathsf{trav}(T(S)))] &= \sum_S p_S E_T[d(\mathsf{trav}(T(S)))] \\ &= \sum_S p_S O(\log n) d(OPT(S)) \\ &= O(\log n) \sum_S p_S d(OPT(S)) \\ &= O(\log n) E_S[d(OPT(S))]. \end{aligned}$$

If $APOPT(S)$ denotes the optimal *a priori* tour restricted to S, then $d(OPT(S)) \leq d(APOPT(S))$ for each S, and so $E_S(d(OPT(S))) \leq E_S(d(APOPT(S)))$, which is exactly the optimal value for the *a priori* TSP. ■

Note however, that this does *not* imply a $O(\log n)$ bound for the universal TSP problem. The universal TSP is concerned with $\max_S d(\mathsf{trav}(T(S))/d(OPT(S))$, and we cannot conclude anything directly about this quantity based on the arguments above. It is interesting to note that we can obtain analogous results for variants of the *a priori* and universal minimum spanning tree and Steiner tree problems as introduced by Bertsimas [4].

Finally, note that the result in [16] implies that it is not possible to find a better approximation for the *a priori* TSP using the technique of probabilistically embedding the metric in *cut metrics* (a cut metric is a metric d that can be represented as $d(i,j) = \sum_{S:|\{i,j\}\cap S|=1} w_S$), of which tree metrics are a subclass.

Acknowledgments

We would like to thank the anonymous referee for pointing out the connection to Kalmanson metrics and the master tour property.

Research of the first author supported partially by NSF grant CCR-0430682, of the second author supported partially by NSF grants CCR-0430682 & DMI-0500263.

References

[1] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996)*, pages 184–193. IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

[2] J. J. Bartholdi, III and L. K. Platzman. An $O(N \log N)$ planar travelling salesman heuristic based on spacefilling curves. *Oper. Res. Lett.*, 1(4):121–125, 1981/82.

[3] J. J. Bartholdi, III, L. K. Platzman, R. L. Collins, and W. H. Warden, III. A minimal technology routing system for meals on wheels. *Interfaces*, 13(3):1–8, 1984.

[4] D. Bertsimas. *Probabilistic Combinatorial Optimization Problems*. PhD thesis, MIT, Cambridge, Mass., 1988.

[5] D. Bertsimas and M. Grigni. Worst-case examples for the spacefilling curve heuristic for the euclidean traveling salesman problem. *Operations Research Letters*, 8(5):241–244, October 1989.

[6] D. J. Bertsimas, P. Jaillet, and A. R. Odoni. A priori optimization. *Operations Research*, 38(6):1019–1033, 1990.

[7] G. Christopher, M. Farach, and M. Trick. The structure of circular decomposable metrics. In *Algorithms—ESA '96 (Barcelona)*, volume 1136 of *Lecture Notes in Comput. Sci.*, pages 486–500. Springer, Berlin, 1996.

[8] V. G. Deĭneko, R. Rudolf, and G. J. Woeginger. Sometimes travelling is easy: the master tour problem. *SIAM J. Discrete Math.*, 11(1):81–93, 1998.

[9] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 448–455 (electronic), New York, 2003. ACM.

[10] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. System Sci.*, 69(3):485–497, 2004.

[11] A. Gupta, M. T. Hajiaghayi, and H. Räcke. Oblivious network design. In *SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 970–979, New York, NY, USA, 2006. ACM Press.

[12] M. T. Hajiaghayi, R. Kleinberg, and T. Leighton. Improved lower and upper bounds for universal tsp in planar metrics. In *SODA '06: Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 649–658, New York, NY, USA, 2006. ACM Press.

[13] P. Jaillet. Probabilistic traveling salesman problems. Technical Report 185, Operations Research Center, MIT, 1985.

[14] P. Jaillet. A priori solution of a traveling salesman problem in which a random subset of the customers are visited. *Operations Research*, 36:929–936, 1988.

[15] L. Jia, G. Lin, G. Noubir, R. Rajaraman, and R. Sundaram. Universal approximations for TSP, Steiner tree, and set cover. In *STOC '05: Proceedings of the Thirty-seventh Annual ACM Symposium on Theory of Computing*, pages 386–395, New York, NY, USA, 2005. ACM Press.

[16] T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM*, 46(6):787–832, 1999.

[17] L. K. Platzman and I. John J. Bartholdi. Spacefilling curves and the planar travelling salesman problem. *J. ACM*, 36(4):719–737, 1989.