On the Complexity of the Metric TSP under Stability Considerations

Matúš Mihalák, Marcel Schöngens, Rastislav Šrámek, and Peter Widmayer
Department of Computer Science, ETH Zurich, Zurich, Switzerland

Abstract. We consider the metric Traveling Salesman Problem (\(\Delta\)-TSP for short) and study how stability (as defined by Bilu and Linial [3]) influences the complexity of the problem. On an intuitive level, an instance of \(\Delta\)-TSP is \(\gamma\)-stable (\(\gamma > 1\)), if there is a unique optimum Hamiltonian tour and any perturbation of arbitrary edge weights by at most \(\gamma\) does not change the edge set of the optimal solution (i.e., there is a significant gap between the optimum tour and all other tours). We show that for \(\gamma \geq 1.8\) a simple greedy algorithm (resembling Prim’s algorithm for constructing a minimum spanning tree) computes the optimum Hamiltonian tour for every \(\gamma\)-stable instance of the \(\Delta\)-TSP, whereas a simple local search algorithm can fail to find the optimum even if \(\gamma\) is arbitrary. We further show that there are \(\gamma\)-stable instances of \(\Delta\)-TSP for every \(1 < \gamma < 2\). These results provide a different view on the hardness of the \(\Delta\)-TSP and give rise to a new class of problem instances which are substantially easier to solve than instances of the general \(\Delta\)-TSP.

1 Introduction

NP-hardness is a common concept of quantifying the complexity of an optimization problem. It can be seen as quite a pessimistic approach, since it considers the worst case running time of an algorithm for all instances of a problem. Hence, many other views on the difficulty of algorithmic problems exist. Specifically, Bilu and Linial [3] observed that an optimization problem that is NP-hard in general may turn out to be easy (i.e., polynomial time solvable) if there is one optimum solution that stands out. The notion of stability captures this idea: If the solution to a combinatorial optimization problem does not change even if we multiply the input parameters by a given factor, we call the problem instance stable with respect to this factor. Stability models a variety of practical considerations, such as measurement errors in input parameters. In this paper we study the well-known Traveling Salesman Problem (TSP) in the context of stability. We limit ourselves to the case in which edge weights satisfy the triangle inequality, the NP-hard metric TSP (\(\Delta\)-TSP).

Our Results. We show a tight upper bound of 2 on the stability of any \(\Delta\)-TSP instance. We prove that any 1.8-stable instance of \(\Delta\)-TSP can be solved in polynomial time by a greedy algorithm, but we provide instances of stability 5/3 on which the same algorithm fails. In the end, we provide a class of Euclidean
instances that are 2-stable and we show that on these instances a simple local search algorithm fails. This result also holds for non-Euclidean instances of arbitrary stability.

Related Work. Bilu and Linial [3] considered the MAX-CUT Problem and showed that $\gamma$-stable instances can be solved correctly in polynomial time on (i) simple graphs of minimum degree $\delta$, when $\gamma > 2n/\delta$, where $n$ is the number of vertices and (ii) weighted graphs of maximal degree $\zeta$, when $\gamma > \sqrt{\zeta n}$. Balcan et al. [2] studied clusterings and restricted themselves to instances that have the $(c, \epsilon)$-property, i.e., instances where any $c$-approximation of the given objective function (e.g. k-median, k-means, etc.) is $\epsilon$-close to the optimal clustering; two clusterings are considered $\epsilon$-close if they differ only in an $\epsilon$-fraction of points. They showed that for such instances one can produce $\epsilon$-close clusterings in polynomial time, even for values of $c$ where a $c$-approximation is provably NP-hard. Awasthi et al. [1] proved that for center-based clustering objectives a constant stability (as defined by Bilu and Linial [3]) is sufficient to obtain an optimal clustering in polynomial time. Further, they relaxed the requirements of the $(c, \epsilon)$-property and showed that one can still find optimal or near optimal clusterings in polynomial time. Spielman and Teng [4] introduced smoothed analysis, which on an intuitive level states that hard input instances occur very rarely at discrete points in solution space and therefore a small perturbation yields a polynomial time solvable input instance. Note that stability focuses on the structure of an instance rather than on the topology of the solution space.

In the following we introduce some notation and provide formal definitions of the required notions. Let $K_n = (V_n, E_n)$ be the complete graph on $n$ vertices where $V_n$ is the vertex set and $E_n$ is the set of all undirected edges on $V_n$. Let $w : E_n \to \mathbb{R}^+$ be a function that assigns a positive weight to each edge. Throughout this paper we assume that $w$ satisfies the triangle inequality, i.e. for any $u, v, x \in V_n$ it holds that $w(u, v) \leq w(u, x) + w(x, v)$. To simplify notation, we may write $w'_e$ instead of $w(e'_e)$, as long as it does not affect readability, and for the weight $w(S) = \sum_{e \in S} w(e)$ of a set $S$ of edges we may write $w_S$. Any sequence of $n$ distinct vertices of $K_n$ defines a Hamiltonian tour $H$. The set of all Hamiltonian tours is denoted by $S(K_n)$. The edge set of $H$ is denoted by $E(H)$ and we may write $w_H$ instead of $w_{E(H)}$. Using the above notation the $\Delta$-TSP is defined as follows. For an input instance $I = (K_n, w)$, find a $\hat{H} \in S(K_n)$ such that for all $H' \in S(K_n)$ we have $w_{\hat{H}} \leq w_{H'}$. We call $\hat{H}$ an optimal solution or optimal Hamiltonian tour of $I$.

We now formalize the notion of stability. When each edge weight of $I$ is multiplied by an individual factor of at least 1 and at most $\gamma$, we say that the instance the instance $I$ is perturbed by at most $\gamma$. Assume that $I$ has a unique optimal Hamiltonian tour $\hat{H}$ (this assumption is justified later). Consider a non-optimal tour $H$ and let $A(H)$ be the set of edges that are part of $H$ but not of $\hat{H}$. Further, let $D(H)$ be the set of edges that are part of $\hat{H}$ but not of $H$. A perturbation causes $\hat{H}$ to become non-optimal iff there is a non-optimal $H$ s.t. the perturbation causes $w_D(H)$ to become greater than $w_{A(H)}$. Thus, there