Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Determination of optimal control strategies for TSP by dynamic programming

Michele Aicardi, Davide Giglio, and Riccardo Minciardi

Abstract— **The traveling salesman problem (TSP) is considered in this paper with the aim of determining optimal control strategies, which provide the optimal decisions as functions of the system state. The adopted TSP model takes into account the travel time between cities and is characterized by the presence of a constraint on the time instant by which a city has to be visited (due date); in this connection, the cost to be minimized is the total weighted tardiness cost, and the decision variables are those concerning the sequence of cities to be visited. The optimal (closed-loop) strategies are determined through a two-step procedure. In the second part of the paper, an extended version of the TSP model, which includes stopover times, is considered, and optimal control strategies are determined also for this model (in this case, through a four-step procedure).**

I. INTRODUCTION

The traveling salesman problem (TSP) [1], [2] is one of the most studied problems in the literature, especially in the field of combinatorial optimization and computation complexity, and this topic continues to attract the attention of researchers from various areas [3]. The original problem statement of TSP is: *"given n cities and their intermediate distance, find a shortest round trip tour that visits each city once and then returns to the starting city"*. Since such statement, several extensions have been proposed in order to solve similar problems. The TSP is known to be NP-hard and the decision version of the problem is NP-complete. Then, researchers have mainly focused their attention to the determination of methods and heuristic algorithms to compute the solution of TSP, in reasonable time [4], [5], [6].

In this paper, the traveling salesman problem is considered under a viewpoint different from the one commonly adopted within combinatorial optimization theory. In fact, the objective of this paper is not the development of an algorithm to find the optimal decisions. Instead, the decision problem is set within a control-theoretic framework, as optimal control strategies are sought, capable of providing, at each decision step, the optimal decisions as functions of the current system state. The advantages of providing optimal control strategies are apparent. Travel times are strongly dependent on the traffic, and then a trip between two cities may last more or less than the "nominal" travel time defined in advance. In addition, other unpredictable events may cause delays on the trip schedule. Then, taking into account the presence of constraints over the visit time instants, the determination of optimal control strategies, instead of optimal decisions, makes the proposed approach utilizable when deviations from the nominal behaviour may be recognized in real-time.

The authors are with the Department of Communications, Computers, and Systems Science (DIST), University of Genova, Via Opera Pia 13, 16145 – Genova, Italy. *Corresponding author:* Davide Giglio, davide.giglio@unige.it

The adopted TSP model takes into account the travel time between cities, and the cost to be minimized in the optimal control problem is the total weighted tardiness cost, where each tardiness is computed by assuming that each city should be visited no later than a certain time instant (due date). The decision variables are binary variables whose values define the sequence of visits in the round trip. An extension of the TSP model is then proposed, where the time intervals of staying in the cities (stopover times) are also considered. Stopover times are not fixed; however, they are constrained between a lower and an upper bound (nominal value). Then, in the extended TSP model, a further decision (continuous) variable exists (stopover time in the next city to be visited) and the cost function takes into account also the deviation of the actual stopover times from the nominal ones.

The optimal control strategies are determined by applying dynamic programming to the considered model. The proposed procedure is of constructive type, hence it does not require the solution of a mixed-integer programming problem. It is worth noting that the proposed methodology is based on some results [7], [8] which have been obtained in connection with the scheduling of jobs, belonging to different classes, on a single machine with controllable processing times [9] and generalized due dates [10].

The rest of the paper is organized as follows. In the next section, the adopted TSP model is presented along with the formalization of the optimal control problem. Section III provides the theorem which allows the determination of optimal control strategies. Section IV outlines an example of the application of the proposed method. The extended version of the TSP model is presented in Section V. Some conclusions are reported in the last section.

II. THE TSP MODEL AND THE OPTIMAL CONTROL PROBLEM

Consider a traveling salesman which must visit K cities, being T_k, $k = 1, \ldots, K$, the generic city to be visited. It is assumed that the salesman starts and finishes its travel in city T_1, and that all the other cities can be visited in whichever sequence (the only precedence constraint is relevant to the initial and final city, i.e., T_1). Moreover, each city (different from T_1) must be visited only once. A trip from one city to another cannot be interrupted (no preemption). The travel time $\xi_{h,k} > 0$ from T_h to T_k, with $h, k = 1, \ldots, K$, $h \neq k$, is fixed and known in advance; $\xi_{h,k}$ is generally allowed to be different from $\xi_{k,h}$. It is further assumed that the

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

stopover times are negligible with respect to the travel time between cities. The salesman starts its travel from T_1 at time 0, and each city should be visited by a certain time instant d_k, $k = 1, \ldots, K$ (d_1 is the scheduled time for returning to T_1 and then for finishing the travel); in case of non fulfillment of such a requirement, a tardiness cost is paid; in this connection, let α_k, $k = 1, \ldots, K$, be the coefficient specifying the unitary tardiness penalty.

This TSP model is similar to the model proposed in [7], [8], where N_k jobs of class P_k, $k = 1, \ldots, K$, have to be executed. In this case, $N_k = 1$, the execution of the jobs corresponds to visit the cities, and due dates correspond to the scheduled times for visiting the cities. Moreover, the setup time $\xi_{h,k}$, $h, k = 1, \ldots, K$, corresponds to the travel time between cities. However, in this definition of TSP model, service times $\tau_{k,i}$, $k = 1, \ldots, K$, $i = 1, \ldots, N_k$, as well as their lower and upper bounds, namely τ_k^{low} and τ_k^{nom}, respectively, are not considered due to the assumption on stopover time. In any case, the results presented in [7], [8] can be applied to the TSP model here proposed.

Consider again the traveling salesman problem; the objective function to be minimized is

$$\sum_{k=1}^{K} \alpha_k \max\{C_k - d_k, 0\} \tag{1}$$

where C_k, $k = 1, \ldots, K$, is the time at which T_k is visited (C_1 is the time at which the salesman returns back home). It is worth noting that, as the cost function to be minimized takes into account the tardiness (and not, for example, the earliness), there is no advantage in delaying any visit if the other ones remain unchanged. Hence, there is an optimal solution of the traveling salesman problem where no idle time is inserted between the visit of two cities.

The aforementioned model can be easily formalized through a state space representation, where the system state, when a new decision has to be taken, i.e., at an instant t at which a city has been visited, is the $(K+2)$-tuple $\underline{x}(t) = [v_1(t), \ldots, v_k(t), h(t), t]^T$, being $v_k(t)$, $k = 1, \ldots, K$, a value equal to 1 if city T_k has been already visited at time instant t, and 0 otherwise, and being $h(t)$ the index of the last visited city.

In this model, the decisions are taken only at specific time instants corresponding to the visit of a city (but for the last one, that is, T_1), and to the initial time instant; thus, decision instants are actually discrete in time (although not equally spaced), and are denoted by t_j, $j = 0, 1, \ldots, K-1$. Since the salesman starts its travel from T_1 at time 0, then $t_0 = 0$. At time instant t_j, j cities have already been visited. The action $\underline{u}(t_j)$, function of $\underline{x}(t_j)$, that has to be taken at time instant t_j corresponds to the choice of the next city to be visited (and then of the next movement). In this connection, let $\delta_k(t_j) \in \{0, 1\}$, $k = 1, \ldots, K$, denote a decision variable whose value is 1 if the next city to be visited is T_k, and 0 otherwise. Obviously, $\sum_{k=1}^{K} \delta_k(t_j) = 1 \; \forall \, t_j$. Thus, $\underline{u}(t_j) = [\delta_1(t_j), \ldots, \delta_K(t_j)]^T$. For the sake of brevity, notations $\underline{x}_j$, $v_{k,j}$, h_j, $\underline{u}_j$, and $\delta_{k,j}$ will be used instead of $\underline{x}(t_j)$, $v_k(t_j)$, $h(t_j)$, $\underline{u}(t_j)$, and $\delta_k(t_j)$, respectively.

Since idle times are not allowed, the state equations of the system can be written as follows

$$\underline{x}_{j+1} = \begin{bmatrix} v_{1,j+1} \\ \vdots \\ v_{K,j+1} \\ h_{j+1} \\ t_{j+1} \end{bmatrix} = \begin{bmatrix} v_{1,j} + \delta_{1,j} \\ \vdots \\ v_{K,j} + \delta_{K,j} \\ \sum_{k=1}^{K} k\delta_{k,j} \\ t_j + \sum_{k=1}^{K} \xi_{h_j,k}\delta_{k,j} \end{bmatrix} \tag{2}$$

$j = 0, 1, \ldots, K-1$, where decision variables $\delta_{k,j}$, $k = 1, \ldots, K$, are constrained by

$$\delta_{1,j} = 0 \qquad \forall \, j = 0, 1, \ldots, K-2 \tag{3}$$

$$\delta_{1,K-1} = 1 \tag{4}$$

$$\sum_{k=1}^{K} \delta_{k,j} = 1 \qquad \forall \, j = 0, 1, \ldots, K-1 \tag{5}$$

$$\sum_{j=0}^{K-1} \delta_{k,j} = 1 \qquad \forall \, k = 1, \ldots, K \tag{6}$$

with initial state $\underline{x}_0 = [0, \ldots, 0, 1, 0]^T$. In figure 1, all possible evolutions of the system state in the case of $K = 4$ are represented. Note that, the state diagram is structured in $K+1$ stages (K decision stages plus the final stage); for each decision stage, a variable number of possible decisions is allowed.

In order to formally state the optimal control problem, it is convenient to rewrite the objective function. To this end, consider the mapping $k(j) = \sum_{l=1}^{K} l \, \delta_{l,j}$, $j = 0, 1, \ldots, K-1$; $k(j)$ is the index of the city which is selected, at time instant t_j, as the next city to be visited (and then $T_{k(j)}$ is to be visited at t_{j+1}). In this way, the objective function (1) can be rewritten as

$$\sum_{j=0}^{K-1} \alpha_{k(j)} \max\{C_{k(j)} - d_{k(j)}, 0\} \tag{7}$$

On the basis of such a notation, it is possible to formalize the following optimal control problem for the aforementioned TSP model.

Problem 1 (Basic TSP): With reference to the dynamic system represented through the state equation (2), and taking into account constraints (3), (4), (5), and (6), find control strategies $\delta_{k,j}^{\circ}(v_{1,j}, \ldots, v_{K,j}, h_j, t_j)$, $k = 1, \ldots, K$, $j = 0, 1, \ldots, K-1$, to be applied at any state $[v_{1,j}, \ldots, v_{K,j}, h_j, t_j]^T$ such that $j = 0, 1, \ldots, K-1$, with t_j non-negative real and $v_{k,j} \in \{0, 1\}$, $k = 1, \ldots, K$, that minimize the objective function (7).

III. OPTIMAL CONTROL STRATEGIES FOR THE BASIC TSP

The main result of the paper (procedure to find the optimal control strategies for the basic TSP) can be obtained by introducing a specific class of functions.

Definition 1: Given an integer number M, $M \geq 1$, and a positive real number ν, a function $f(x)$ is said to be a

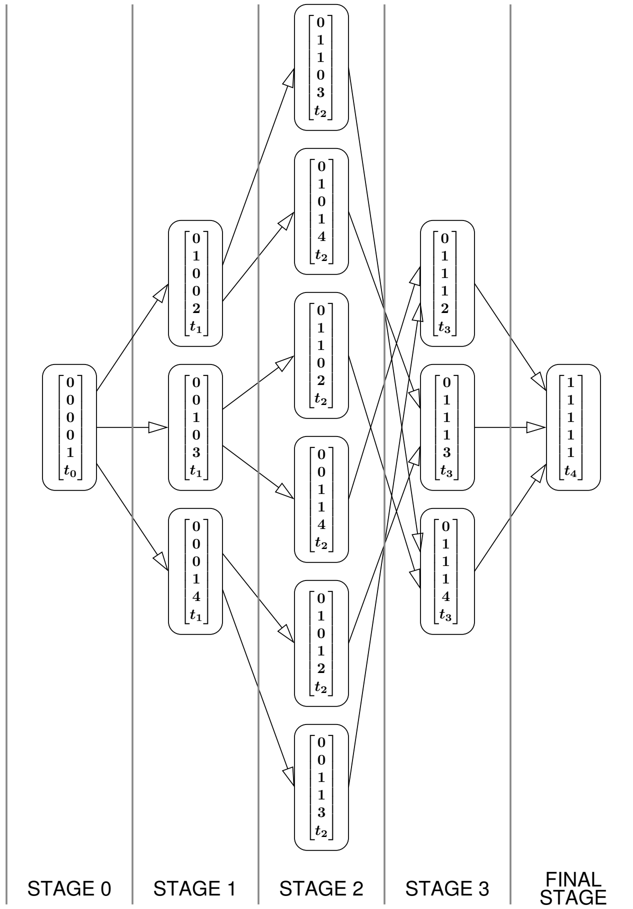

Fig. 1. A representation of the state diagram of the TSP model, when $K = 4$.

"$\mathcal{PWL}(M,\nu)$-function" (Piece-Wise Linear(M,ν)-function) if:

- $f(x)$ is a non-decreasing continuous piece-wise linear function of x, characterized by:
  - M break points γ_i, $i = 1,\dots,M$, being $\gamma_{i+1} > \gamma_i$, $i = 1,\dots,M-1$;
  - $(M+1)$ slopes, namely μ_0 in interval $(-\infty,\gamma_1)$, μ_i in (γ_i,γ_{i+1}), $i = 1,\dots,M-1$, and μ_M in $(\gamma_M,+\infty)$; of course, $\mu_{i+1} \neq \mu_i$, $i = 0,\dots,M-1$;
- $f(x) = 0$ for any $x \le \gamma_1$, that is, $\mu_0 = 0$;
- $\mu_i \ge \nu$, $i = 1,\dots,M$. $\triangle$

The graphical structure of a generic $\mathcal{PWL}(M,\nu)$-function $f(x)$ is represented in figure 2.

Moreover, it is necessary to introduce two preliminary results regarding the sum of two $\mathcal{PWL}(\cdot,\cdot)$-functions and the minimum of k $\mathcal{PWL}(\cdot,\cdot)$-functions. The proofs of the two lemmas are here omitted (they can be found in [8]).

Lemma 1: Let $f(x)$ be a $\mathcal{PWL}(M,\nu)$-function, whose first break point is γ_1^{f}, and $g(x)$ a $\mathcal{PWL}(N,\xi)$-function, whose first break point is γ_1^{g}. Then the function

$$h(x) = f(x) + g(x) \tag{8}$$

is a $\mathcal{PWL}(Q,\rho)$-function, where $\rho = \min\{\nu,\xi\}$ and $1 \le Q \le M+N$, having the first break point in $\min\{\gamma_1^{\mathrm{f}},\gamma_1^{\mathrm{g}}\}$. $\square$

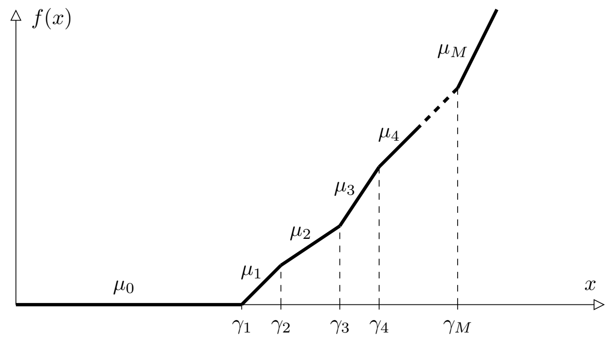

Fig. 2. A $\mathcal{PWL}(M,\nu)$-function $f(x)$.

Lemma 2: Let $f_k(x)$ be a $\mathcal{PWL}(M_k,\nu_k)$-function having its first break point in γ_1^k, $k = 1,\dots,K$ (being K an arbitrary integer number). Then the function

$$h(x) = \min\big\{f_k(x), k = 1,\dots,K\big\} \tag{9}$$

is a $\mathcal{PWL}(Q,\rho)$-function, where $\rho = \min\{\nu_k, k = 1,\dots,K\}$ and $1 \le Q \le 2\big(2^{K-2}\cdot M_1 + (\sum_{i=2}^{K} 2^{K-i}\cdot M_i) - (2^{K-1}-1)\big)$, having the first break point in $\max\{\gamma_1^k, k = 1,\dots,K\}$. $\square$

Consider now Problem 1, and let $\mathcal{J}^{\circ}_{v_1,\dots,v_K,h}(t_{\tilde v})$, where $\tilde v = \sum_{k=1}^{K} v_k$, denote the optimal cost-to-go associated with state $[v_1,\dots,v_K,h,t_{\tilde v}]^T$ (note that $\tilde v = j$). When $\tilde v = K$ (final stage), the optimal cost-to-go is obviously zero, whereas when $\tilde v < K$ it can be expressed as

$$\mathcal{J}^{\circ}_{v_1,\dots,v_K,h}(t_{\tilde v}) = \\ = \min_{\substack{\delta_{1,j},\dots,\delta_{K,j}\\ j=\tilde v,\dots,K}} \sum_{j=\tilde v}^{K-1} \alpha_{k(j)} \max\{C_{k(j)} - d_{k(j)}, 0\} \tag{10}$$

where the following constraints have to be fulfilled in the minimization:

- $\delta_{1,j} = 0$, $\forall\, j = \tilde v,\dots,K-2$;
- $\delta_{1,K-1} = 1$;
- $\sum_{k=1}^{K} \delta_{k,j} = 1$, $\forall\, j = \tilde v,\dots,K-1$;
- $\sum_{j=\tilde v}^{K-1} \delta_{k,j} = 1$, for any state $[v_1,\dots,v_K,h,t_{\tilde v}]^T$ such that $v_k = 0$, $k = 1,\dots,K$, $\forall\, j = \tilde v,\dots,K-1$;

along with the state equation (2). Such constraints will be understood in any expression of the cost-to-go hereafter.

Then, taking into account the absence of idle times in the service sequence, the general dynamic programming recursion for the determination of optimal control strategies for the Basic TSP (Problem 1) is

$$\mathcal{J}^{\circ}_{v_1,\dots,v_K,h}(t_{\tilde v}) = \\ = \min_{\substack{\delta_{k,\tilde v}\\ k=1,\dots,K}} \left\{ \sum_{k=1}^{K} \delta_{k,\tilde v}\Big[\alpha_k \max\{t_{\tilde v} + \xi_{h,k} - d_k, 0\} + \right. \\ \left. + \mathcal{J}^{\circ}_{v_1,\dots,v_k+1,\dots,v_K,k}(t_{\tilde v} + \xi_{h,k})\Big]\right\} \tag{11}$$

being $[v_1, \ldots, v_k+1, \ldots, v_K, k, t_{\tilde{v}}+\xi_{h,k}]^T$ the state reached from $[v_1, \ldots, v_K, h, t_{\tilde{v}}]^T$, when city k is visited after city h. Note that, since $t_{\tilde{v}+1} = t_{\tilde{v}} + \xi_{h,k}$, the optimal cost-to-go in (11), relevant to state $[v_1, \ldots, v_k+1, \ldots, v_K, k, t_{\tilde{v}+1}]^T$, can be expressed as function of the current time instant. Cost-to-go (11) is equivalent to

$$\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}}) = \min\Big\{\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{1,\tilde{v}}=1), \ldots \mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{K,\tilde{v}}=1)\Big\} \quad (12)$$

where symbol $\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}}=1)$ denotes the value of $\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}})$ conditioned to the choice of the control variable $\delta_{k,\tilde{v}}=1$, $k=1,\ldots,K$.

Then, a general result concerning the solution to Problem 1 can be expressed in the following form.

Theorem 1: The optimal control strategies solving Problem 1 (Basic TSP) can be obtained through a two-steps procedure.

1) Determine, for each $(K+1)$-tuple $(v_1, \ldots, v_K, h)$, such that $v_k \in \{0,1\}$, $k=1,\ldots,K$, $\tilde{v} < K$, and $h \in \{1,\ldots,K\}$, the conditioned costs-to-go and the optimal costs-to-go through the backward recursive relations

$$\begin{aligned}\mathcal{J}_{v_1,\ldots,v_K,h}&(t_{\tilde{v}} \mid \delta_{k,\tilde{v}}=1) = \\ &= \alpha_k \max\{t_{\tilde{v}} + \xi_{h,k} - d_k, 0\} + \\ &\quad + \mathcal{J}^\circ_{v_1,\ldots,v_k+1,\ldots,v_K,k}(t_{\tilde{v}} + \xi_{h,k})\end{aligned} \quad (13)$$

for each $k \in \mathcal{A}_{v_1,\ldots,v_K,h}$, being $\mathcal{A}_{v_1,\ldots,v_K,h} = \{i : i \in \{2,\ldots,K\}$ and $v_i = 0\}$ when $\tilde{v} < K-1$ and $\mathcal{A}_{v_1,\ldots,v_K,h} = \{1\}$ when $\tilde{v} = K-1$,

$$\begin{aligned}&\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}}) = \\ &= \min_{k \in \mathcal{A}_{v_1,\ldots,v_K,h}} \big\{\mathcal{J}_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}}=1)\big\}\end{aligned} \quad (14)$$

Observe that $\mathcal{A}_{v_1,\ldots,v_K,k}$ is the set of indexes of cities that still have to be visited in state $[v_1, \ldots, v_K, k, t_{\tilde{v}+1}]^T$. Backward recursion (13) is initialized by

$$\mathcal{J}^\circ_{1,\ldots,1,1}(t_K) = 0 \quad (15)$$

2) Then, the optimal control strategies, for each $(K+1)$-tuple $(v_1, \ldots, v_K, h)$, such that $v_k \in \{0,1\}$, $k = 1,\ldots,K$, and $\tilde{v} < K$, are obtained as
   - if $\tilde{v} \le K-3$:

$$\begin{aligned}&\delta^\circ_{k,\tilde{v}}(v_1,\ldots,v_K,h,t_{\tilde{v}}) = \\ &= \begin{cases} 1 & \text{if } k = \operatorname{argmin}_p \big\{ \\ & \quad \mathcal{J}_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{p,\tilde{v}}=1), \\ & \quad p \in \mathcal{A}_{v_1,\ldots,v_K,h}\big\} \\ 0 & \text{otherwise}\end{cases}\end{aligned} \quad (16)$$

   (in case of multiple values of the argmin, ties are broken arbitrarily)
   - if $\tilde{v} = K-2$:

$$\begin{aligned}&\delta^\circ_{k,\tilde{v}}(v_1,\ldots,v_K,h,t_{\tilde{v}}) = \\ &= \begin{cases} 1 & \text{if } k \in \{2,\ldots,K\} \text{ and } v_k = 0 \\ 0 & \text{otherwise}\end{cases}\end{aligned} \quad (17)$$

   - if $\tilde{v} = K-1$:

$$\delta^\circ_{k,\tilde{v}}(v_1,\ldots,v_K,h,t_{\tilde{v}}) = \begin{cases} 1 & \text{if } k=1 \\ 0 & \text{otherwise}\end{cases} \quad (18)$$

Besides, the optimal cost-to-go $\mathcal{J}^\circ_{v_1,\ldots,v_K,h}(t_{\tilde{v}})$ is a $\mathcal{PWL}(M, \hat{\alpha})$-function, for some M and where $\hat{\alpha} = \min\{\alpha_1, \ldots, \alpha_K\}$, for any set $(v_1, \ldots, v_K, h)$ such that $v_k \in \{0,1\}$, $k = 1,\ldots,K$, and $\tilde{v} < K$. □

Proof: The proof of Theorem 1 is based on the application of dynamic programming. Consider the system state diagram (illustrated in figure 1) and the first decision stage, that is, stage $K-1$.

In the considered TSP model, the decision to be taken in this stage, at time instant t_{K-1}, is mandatory, since all cities but T_1 have been already visited, and then the traveling salesman must return to its starting point. Then, the optimal control strategies are

$$\delta_{1,K-1}(0,1,\ldots,1,h,t_{K-1}) = 1 \quad (19)$$

$$\delta_{k,K-1}(0,1,\ldots,1,h,t_{K-1}) = 0 \quad (20)$$

$\forall\, h,k \in \{2,\ldots,K\}$, being $[0,1,\ldots,1,h,t_{K-1}]^T$ the generic state of stage $K-1$. Then, taking into account (15), the optimal cost-to-go of the generic state in stage $K-1$ is simply

$$\begin{aligned}\mathcal{J}^\circ_{0,1,\ldots,1,1,h}(t_{K-1}) &= \mathcal{J}_{0,1,\ldots,1,1,h}(t_{K-1} \mid \delta_{1,K-1}=1) = \\ &= \alpha_1 \max\{t_{K-1} + \xi_{h,1} - d_1, 0\}\end{aligned} \quad (21)$$

for each $h = 2,\ldots,K$. Such optimal cost-to-go is a $\mathcal{PWL}(1,\alpha_1)$-function.

Consider now stage $K-2$. The decision to be taken at t_{K-2}, in the states belonging to this stage, is mandatory again. The fact is that the generic state of this stage, namely $[0,1,\ldots,0,\ldots,1,h,t_{K-2}]^T$, is characterized by $v_1 = 0$, $v_j = 0$, and $v_k = 1$, for some $j \in \{2,\ldots,K\}$ and for any $k \in \{2,\ldots,K\}$, $k \ne j$. Then, the next city to be visited is T_j, since T_1 must be the last city to be visited. Thus, the optimal control strategies are

$$\delta_{j,K-2}(0,1,\ldots,0,\ldots,1,h,t_{K-2}) = 1 \quad (22)$$

$$\delta_{k,K-2}(0,1,\ldots,0,\ldots,1,h,t_{K-2}) = 0 \quad (23)$$

for any $k \in \{2,\ldots,K\}$, $k \ne j$. The optimal cost-to-go in the generic state in stage $K-2$ is simply

$$\begin{aligned}&\mathcal{J}^\circ_{0,1,\ldots,0,\ldots,1,h}(t_{K-2}) = \\ &\quad = \mathcal{J}_{0,1,\ldots,0,\ldots,1,h}(t_{K-2} \mid \delta_{j,K-2}=1) = \\ &\quad = \alpha_j \max\{t_{K-2} + \xi_{h,j} - d_j, 0\} + \mathcal{J}^\circ_{0,1,\ldots,1,j}(t_{K-1})\end{aligned} \quad (24)$$

Note that the optimal cost-to-go $\mathcal{J}^\circ_{0,1,\ldots,1,j}(t_{K-1})$ can be expressed as function of t_{K-2}, as $t_{K-1} = t_{K-2} + \xi_{h,j}$; then, $\mathcal{J}^\circ_{0,1,\ldots,1,j}(t_{K-2}+\xi_{h,j})$ is still a $\mathcal{PWL}(1,\alpha_1)$-function. Also $\alpha_j \max\{t_{K-2}+\xi_{h,j}-d_j, 0\}$ is a $\mathcal{PWL}(1,\alpha_j)$-function and then, by applying Lemma 1 to the r.h.s. of (24), it is possible to conclude that the optimal cost-to-go of the generic state of stage $K-2$ is a $\mathcal{PWL}(Z_{0,1,\dots,0,\dots,1,h}, \hat{\alpha})$-function, where $Z_{0,1,\dots,0,\dots,1,h} \le 2$ and $\hat{\alpha} = \min\{\alpha_1, \dots, \alpha_K\}$.

Consider now stage $K-3$ and its generic state $[0,1,\dots,0,\dots,0,\dots,1,h,t_{K-2}]^T$, which is characterized by $v_1 = 0$, $v_i = 0$, $v_j = 0$, for some $i,j \in \{2,\dots,K\}$, $i \ne j$, and $v_k = 1$, for any $k \in \{2,\dots,K\}$, $k \ne i,j$. At this stage, a decision has to be taken since it is necessary to choice the next city to be visited among T_i and T_j. The optimal cost-to-go is

$$\begin{aligned}&\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3}) = \\ &= \min_{\delta_{i,K-3},\delta_{j,K-3}} \Big\{ \delta_{i,K-3}\Big[\alpha_i \max\{t_{K-3}+\xi_{h,i}-d_i,0\}+ \\ &\quad + \mathcal{J}^\circ_{0,1,\dots,0,\dots,1,i}(t_{K-2})\Big] + \\ &\quad + \delta_{j,K-3}\Big[\alpha_j \max\{t_{K-3}+\xi_{h,j}-d_i,0\}+ \\ &\quad + \mathcal{J}^\circ_{0,1,\dots,0,\dots,1,j}(t_{K-2})\Big]\Big\}\end{aligned} \tag{25}$$

The minimization in (25) has to be carried out with respect to the binary decision variables $\delta_{i,K-3}$ and $\delta_{j,K-3}$. Consider first the case in which $\delta_{i,K-3} = 1$ (hence $\delta_{j,K-3} = 0$). In this case, the conditioned cost-to-go is

$$\begin{aligned}&\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{i,K-3}=1) = \\ &= \alpha_i \max\{t_{K-3}+\xi_{h,i}-d_i,0\} + \mathcal{J}^\circ_{0,1,\dots,0,\dots,1,i}(t_{K-2})\end{aligned} \tag{26}$$

The optimal cost-to-go $\mathcal{J}^\circ_{0,1,\dots,0,\dots,1,i}(t_{K-2})$ can be expressed as a function of t_{K-3}, as $t_{K-2} = t_{K-3} + \xi_{h,i}$, and then, the r.h.s of (26) is the sum of a $\mathcal{PWL}(1,\alpha_i)$-function with a $\mathcal{PWL}(Z_{0,1,\dots,0,\dots,1,i}, \hat{\alpha})$-function. By applying Lemma 1, it turns out that the conditioned cost-to-go $\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{i,K-3} = 1)$ is a $\mathcal{PWL}(Y^i_{0,1,\dots,0,\dots,0,\dots,1,h}, \hat{\alpha})$-function, where $Y^i_{0,1,\dots,0,\dots,0,\dots,1,h} \le 3$.

In the same way, when $\delta_{j,K-3} = 1$ (hence $\delta_{i,K-3} = 0$), the conditioned cost-to-go is

$$\begin{aligned}&\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{j,K-3}=1) = \\ &= \alpha_j \max\{t_{K-3}+\xi_{h,j}-d_j,0\} + \mathcal{J}^\circ_{0,1,\dots,0,\dots,1,j}(t_{K-2})\end{aligned} \tag{27}$$

which is a $\mathcal{PWL}(Y^j_{0,1,\dots,0,\dots,0,\dots,1,h}, \hat{\alpha})$-function, where $Y^j_{0,1,\dots,0,\dots,0,\dots,1,h} \le 3$.

Having considered separately the two cases, it is possible to determine the optimal cost-to-go $\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3})$, which is simply

$$\begin{aligned}&\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3}) = \\ &= \min\Big\{\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{i,K-3}=1), \\ &\quad , \mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{j,K-3}=1)\Big\}\end{aligned} \tag{28}$$

and, owing to Lemma 2, $\mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3})$ turns out to be a $\mathcal{PWL}(Z_{0,1,\dots,0,\dots,0,\dots,1,h}, \hat{\alpha})$-function, where $Z_{0,1,\dots,0,\dots,0,\dots,1,h} \le 2(Y^i_{0,1,\dots,0,\dots,0,\dots,1,h} + Y^j_{0,1,\dots,0,\dots,0,\dots,1,h} - 1)$.

Then, the optimal control strategies at the generic state $[0,1,\dots,0,\dots,0,\dots,1,h,t_{K-2}]^T$ of stage $K-3$, with $v_1 = 0$, $v_i = 0$, $v_j = 0$, for some $i,j \in \{2,\dots,K\}$, $i \ne j$, and $v_k = 1$, for any $k \in \{2,\dots,K\}$, $k \ne i,j$, are given by

$$\begin{aligned}&\delta^\circ_{i,K-3}(0,1,\dots,0,\dots,0,\dots,1,h,t_{K-3}) = \\ &= \begin{cases} 1 & \text{if } \mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{i,K-3}=1) \le \\ & \quad \le \mathcal{J}^\circ_{0,1,\dots,0,\dots,0,\dots,1,h}(t_{K-3} \mid \delta_{j,K-3}=1) \\ 0 & \text{otherwise} \end{cases}\end{aligned} \tag{29}$$

$$\begin{aligned}&\delta^\circ_{j,K-3}(0,1,\dots,0,\dots,0,\dots,1,h,t_{K-3}) = \\ &= 1 - \delta^\circ_{i,K-3}(0,1,\dots,0,\dots,0,\dots,1,h,t_{K-3})\end{aligned} \tag{30}$$

$$\delta^\circ_{k,K-3}(0,1,\dots,0,\dots,0,\dots,1,h,t_{K-3}) = 0 \tag{31}$$

for any $k \in \{2,\dots,K\}$, $k \ne i,j$.

At this point, by the induction principle, it is only necessary to prove that the following implication is true: "assume that the cost-to-go relevant to states belonging to stage $(\tilde{v}+1)$ and reachable from $[v_1,\dots,v_K,h,t_{\tilde{v}}]^T$, namely $\mathcal{J}^\circ_{v_1,\dots,v_k+1,\dots,v_K,k}(t_{\tilde{v}+1})$, $k \in \mathcal{A}_{v_1,\dots,v_K,h}$, are $\mathcal{PWL}(Z_{v_1,\dots,v_k+1,\dots,v_K,k}, \hat{\alpha})$-functions; then, the optimal control strategies at state $[v_1,\dots,v_K,h,t_{\tilde{v}}]^T$ are those provided by (16), and the optimal cost-to-go $\mathcal{J}^\circ_{v_1,\dots,v_K,h}(t_{\tilde{v}})$ is a $\mathcal{PWL}(M, \hat{\alpha})$-function, for some M".

To prove such implication, consider the decision to be taken in state $[v_1,\dots,v_K,h,t_{\tilde{v}}]^T$. If decision $\delta_{k,\tilde{v}} = 1$ is taken (assuming $k \in \mathcal{A}_{v_1,\dots,v_K,h}$), then the cost-to-go (conditioned) is

$$\begin{aligned}&\mathcal{J}^\circ_{v_1,\dots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}}=1) = \\ &= \alpha_k \max\{t_{\tilde{v}}+\xi_{h,k}-d_k,0\} + \mathcal{J}^\circ_{v_1,\dots,v_k+1,\dots,v_K,k}(t_{\tilde{v}+1})\end{aligned} \tag{32}$$

As $t_{\tilde{v}+1} = t_{\tilde{v}} + \xi_{h,k}$, the optimal cost-to-go $\mathcal{J}^\circ_{v_1,\dots,v_k+1,\dots,v_K,k}(t_{\tilde{v}+1})$ is a $\mathcal{PWL}(Z_{v_1,\dots,v_k+1,\dots,v_K,k}, \hat{\alpha})$-function of $t_{\tilde{v}}$. Then, by applying Lemma 1, it turns out that the conditioned cost-to-go $\mathcal{J}^\circ_{v_1,\dots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}} = 1)$ is a $\mathcal{PWL}(Y^k_{v_1,\dots,v_K,h}, \hat{\alpha})$-function, where $Y^k_{v_1,\dots,v_K,h} \le Z_{v_1,\dots,v_k+1,\dots,v_K,k} + 1$.

Once determined the conditioned costs-to-go for all values of $k \in \mathcal{A}_{v_1,\dots,v_K,h}$, the optimal cost-to-go is simply

$$\begin{aligned}&\mathcal{J}^\circ_{v_1,\dots,v_K,h}(t_{\tilde{v}}) = \\ &= \min_{k \in \mathcal{A}_{v_1,\dots,v_K,h}} \big\{\mathcal{J}_{v_1,\dots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}}=1)\big\}\end{aligned} \tag{33}$$

and, then, the optimal control strategies are

$$\begin{aligned}&\delta^\circ_{k,\tilde{v}}(v_1,\dots,v_K,h,t_{\tilde{v}}) = \\ &= \begin{cases} 1 & \text{if } k = \operatorname{argmin}_p \big\{ \\ & \quad \mathcal{J}_{v_1,\dots,v_K,h}(t_{\tilde{v}} \mid \delta_{p,\tilde{v}}=1), \\ & \quad p \in \mathcal{A}_{v_1,\dots,v_K,h}\big\} \\ 0 & \text{otherwise} \end{cases}\end{aligned} \tag{34}$$

Moreover, owing to Lemma 2, $\mathcal{J}^\circ_{v_1,\dots,v_K,h}(t_{\tilde{v}})$ turns out to be a $\mathcal{PWL}(Z_{v_1,\dots,v_K,h}, \hat{\alpha})$-function, where $Z_{v_1,\dots,v_K,h} \le 2\big(2^{K-2}Y^k_{v_1,\dots,v_K,h} + (\sum_{i=2}^K 2^{K-i}Y^i_{v_1,\dots,v_K,h}) - (2^{K-1}-1)\big)$. This complete the proof. ∎

IV. EXAMPLE

A traveling salesman must visit Alessandria, Milano, Piacenza, and Torino, starting from Genova, in the northern part of Italy (see figure 3). The proposed TSP model, with $K = 5$ is adopted in order to find optimal control strategies.

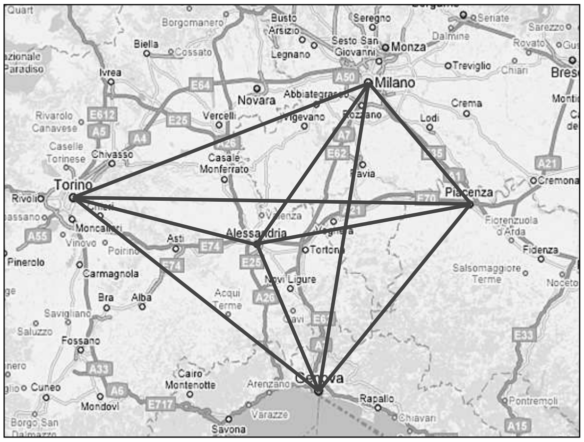

Fig. 3. Example – Map of the considered area.

The travel salesman starts from Genova at 8:00 a.m. and should reach Alessandria by 12:15 p.m., Milano by 11:00 a.m. (with higher priority), Piacenza by 10:30 a.m., and Torino by 11:15 a.m. Moreover, it should return to Genova by 1:00 p.m. Then, due dates are those reported in table I, where unitary tardiness penalty coefficients are also indicated. In table II, travel times (minutes) between cities are reported.

TABLE I
DUE DATES AND UNITARY TARDINESS PENALTY COEFFICIENTS.

City		Due date		Coefficient	
T_1	Genova	d_1	300	α_1	1
T_2	Alessandria	d_2	255	α_2	1
T_3	Milano	d_3	180	α_3	2
T_4	Piacenza	d_4	150	α_4	1
T_5	Torino	d_5	195	α_5	1

TABLE II
TRAVEL TIMES BETWEEN CITIES (MINUTES).

	T_1	T_2	T_3	T_4	T_5
T_1	0	63	99	99	111
T_2	69	0	73	70	70
T_3	104	70	0	61	114
T_4	108	68	62	0	115
T_5	121	73	108	119	0

Theorem 1 can be applied to find the optimal control strategies for each of the 34 states of the state space. As an example, the optimal control strategies in initial state $[0, 0, 0, 0, 0, 1, t_0]^T$, namely $\delta^\circ_{k,0}(0, 0, 0, 0, 0, 1, t_0)$, $k = 1, \dots, 5$, are

$$\delta^\circ_{1,0}(0, 0, 0, 0, 0, 1, t_0) = 0 \tag{35}$$

$$\delta^\circ_{2,0}(0, 0, 0, 0, 0, 1, t_0) = \begin{cases} 1 & \text{if } t_0 \geq 166 \\ 0 & \text{if } t_0 < 166 \end{cases} \tag{36}$$

$$\delta^\circ_{3,0}(0, 0, 0, 0, 0, 1, t_0) = \\ = \begin{cases} 1 & \text{if } -68 \leq t_0 \leq -4 \cup 39 \leq t_0 \leq 166 \\ 0 & \text{if } t_0 < -68 \cup -4 \leq t_0 \leq 39 \cup t_0 \geq 166 \end{cases} \tag{37}$$

$$\delta^\circ_{4,0}(0, 0, 0, 0, 0, 1, t_0) = \\ = \begin{cases} 1 & \text{if } t_0 < -68 \cup -4 \leq t_0 \leq 39 \\ 0 & \text{if } -68 \leq t_0 \leq -4 \cup t_0 \geq 39 \end{cases} \tag{38}$$

$$\delta^\circ_{5,0}(0, 0, 0, 0, 0, 1, t_0) = 0 \tag{39}$$

Such control strategies can be summarized as in figure 4, where the next (first) city to be visited is expressed as function of the initial time instant t_0. Then, in case of start at 8:00 a.m., as scheduled, the first city to be visited is T_4 (Piacenza); however, in case of a 1-hour delay, the first city to be visited becomes T_3 (Milano); T_2 (Alessandria) is the first city to be visited only in the case of severe delays; note also that T_1 and T_5 cannot be the first city to be visited in any case (in particular, T_1 is always the last city to be visited being the starting / ending point).

It is worth noting that also the negative values of t_0 have been considered. In fact, the traveling salesman could start its travel before the scheduled time. In this case, a different optimal sequence of visits could arise. As an example, if the salesman starts at 7:00 a.m., then the first city to be visited is T_3 (Milano) and not T_4 (Piacenza).

V. THE EXTENDED TSP MODEL

Assume now that the stopover times are not negligible with respect to the travel times between cities. In this connection, let s_k, $k = 1, \dots, K$, be the stopover time in city T_k. s_k is a continuous variable whose value ranges from a lower bound s_k^{L} up to an upper bound s_k^{U}, which is also the nominal ("standard") value of such stopover time. The stopover time can be reduced in order, for example, to cope with urgent due dates.

The reduction of the stopover time may be attained, in general, at the price of the payment of an extra cost. Then, the objective function to be minimized becomes

$$\sum_{k=1}^{K} \left[\alpha_k \max\{C_k - d_k, 0\} + \beta_k (s_k^{\mathrm{U}} - s_k) \right] \tag{40}$$

where β_k is a weighting coefficients. It is assumed that $\beta_k = \hat{\beta}$ and that $\alpha_k > \hat{\beta}$, for any $k = 1, \dots, K$. The first assumption states that the extra-cost paid for the reduction of the stopover time is simply proportional to this reduction, with a coefficient that isn't dependent on the city; the second assumption states that any unitary tardiness cost is greater than the unitary cost related to the deviation from the nominal stopover time.

The stopover time is a continuous decision variable. In this connection, let τ_j, $j = 0, 1, \dots, K-1$, indicate the time, determined at t_j, which represents the stopover time for the next city to be visited (that is, for the $(j+1)$-th city).

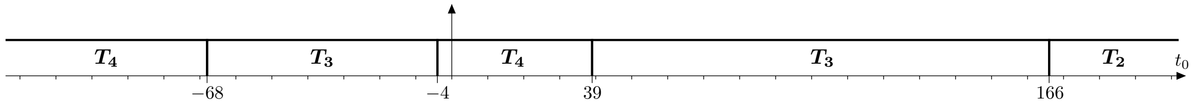

Fig. 4. Example – First city to be visited expressed as function of initial time instant t_0.

The state equations of the system are now

$$\underline{x}_{j+1} = \begin{bmatrix} v_{1,j+1} \\ \vdots \\ v_{K,j+1} \\ h_{j+1} \\ t_{j+1} \end{bmatrix} = \begin{bmatrix} v_{1,j} + \delta_{1,j} \\ \vdots \\ v_{K,j} + \delta_{K,j} \\ \sum_{k=1}^{K} k\delta_{k,j} \\ t_j + \sum_{k=1}^{K} \xi_{h_j,k}\delta_{k,j} + \tau_j \end{bmatrix} \tag{41}$$

$j = 0, 1, \ldots, K-1$, where $\delta_{k,j}$, $k = 1, \ldots, K$, are constrained by (3)–(6), and where τ_j is constrained by

$$\sum_{k=1}^{K} (s_k^{\mathrm{L}} \delta_{k,j}) \le \tau_j \le \sum_{k=1}^{K} (s_k^{\mathrm{U}} \delta_{k,j}) \quad \forall\, j = 0, 1, \ldots, K-1 \tag{42}$$

It is further assumed that $s_1^{\mathrm{L}} = s_1^{\mathrm{U}} = 0$, which means that

$$\tau_{K-1} = 0 \tag{43}$$

The objective function (40) can be rewritten as

$$\sum_{j=0}^{K-1} \left[\alpha_{k(j)} \max\{C_{k(j)} - d_{k(j)}, 0\} + \hat{\beta}(s_{k(j)}^{\mathrm{U}} - \tau_j) \right] \tag{44}$$

and, then, it is possible to consider the following optimization problem for the extended TSP model.

Problem 2 (Extended TSP): With reference to the dynamic system represented through the state equation (41), and taking into account constraints (3), (4), (5), (6), (42), and (43), find control strategies $\delta_{k,j}^{\circ}(v_{1,j}, \ldots, v_{K,j}, h_j, t_j)$, $k = 1, \ldots, K$, $j = 0, 1, \ldots, K-1$, and $\tau_j^{\circ}(v_{1,j}, \ldots, v_{K,j}, h_j, t_j)$, $j = 0, 1, \ldots, K-1$, to be applied at any state $[v_{1,j}, \ldots, v_{K,j}, h_j, t_j]^T$ such that $j = 0, 1, \ldots, K-1$, with t_j non-negative real and $v_{k,j} \in \{0, 1\}$, $k = 1, \ldots, K$, that minimize the objective function (44).

The optimal cost-to-go which is defined for the general dynamic programming recursion is now

$$\begin{aligned} &\mathcal{J}_{v_1,\ldots,v_K,h}^{\circ}(t_{\tilde{v}}) = \\ &= \min_{\substack{\delta_{k,\tilde{v}},\tau_{\tilde{v}} \\ k=1,\ldots,K}} \Bigg\{ \sum_{k=1}^{K} \delta_{k,\tilde{v}} \Big[\alpha_k \max\{t_{\tilde{v}} + \xi_{h,k} + \tau_{\tilde{v}} - d_k, 0\} + \\ &\quad + \hat{\beta}(s_k^{\mathrm{U}} - \tau_{\tilde{v}}) + \mathcal{J}_{v_1,\ldots,v_k+1,\ldots,v_K,k}^{\circ}(t_{\tilde{v}} + \xi_{h,k} + \tau_{\tilde{v}}) \Big] \Bigg\} \end{aligned} \tag{45}$$

Then, the solution to Problem 2 is provided by the following theorem (whose proof is here omitted for the sake of brevity).

Theorem 2: The optimal control strategies solving Problem 2 (Extended TSP) can be obtained through a four-steps procedure.

1) Determine, for each (K+1)-tuple $(v_1, \ldots, v_K, h)$, such that $v_k \in \{0, 1\}$, $k = 1, \ldots, K$, $\tilde{v} < K$, and $h \in \{1, \ldots, K\}$, the set of coefficients $\lambda_{v_1,\ldots,v_K,h}^{k}$, $k \in \mathcal{A}_{v_1,\ldots,v_K,h}$, through a backward recursion expressed by

$$\lambda_{v_1,\ldots,v_K,h}^{k} = \min\Big\{ dd_k, \max_{i \in \mathcal{A}_{v_1,\ldots,v_k+1,\ldots,v_K,k}} \Big\{ \lambda_{v_1,\ldots,v_k+1,\ldots,v_K,k}^{i} - s_i^{\mathrm{U}} \Big\} \Big\} - \xi_{h,k} \tag{46}$$

Note that $\mathcal{A}_{v_1,\ldots,v_k+1,\ldots,v_K,k}$ is the set of indexes of cities that still have to be visited in state $[v_1, \ldots, v_k + 1, \ldots, v_K, k, t_{\tilde{v}+1}]^T$, thus indicating also the states reachable from $[v_1, \ldots, v_k+1, \ldots, v_K, k, t_{\tilde{v}+1}]^T$. Note that, in correspondence of final state $[1, \ldots, 1, 1, t_K]^T$, the set $\mathcal{A}_{1,\ldots,1,1}$ is an empty set; this observation provides the way how recursion (46) is initialized.

2) Build, for each (K+1)-tuple $(v_1, \ldots, v_K, h)$, such that $v_k \in \{0, 1\}$, $k = 1, \ldots, K$, $\tilde{v} < K$, and $h \in \{1, \ldots, K\}$, the set of functions $\tau_{\tilde{v}}^{k\,\circ}(v_1, \ldots, v_K, h, t_{\tilde{v}})$, $k \in \mathcal{A}_{v_1,\ldots,v_K,h}$, as follows

$$\begin{aligned} &\tau_{\tilde{v}}^{k\,\circ}(v_1, \ldots, v_K, h, t_{\tilde{v}}) = \\ &= \begin{cases} s_k^{\mathrm{U}} & t_{\tilde{v}} \le \lambda_{v_1,\ldots,v_K,h}^{k} - s_k^{\mathrm{U}} \\ -t_{\tilde{v}} + \lambda_{v_1,\ldots,v_K,h}^{k} & \lambda_{v_1,\ldots,v_K,h}^{k} - s_k^{\mathrm{U}} < t_{\tilde{v}} < \\ & \qquad < \lambda_{v_1,\ldots,v_K,h}^{k} - s_k^{\mathrm{L}} \\ s_k^{\mathrm{L}} & t_{\tilde{v}} \ge \lambda_{v_1,\ldots,v_K,h}^{k} - s_k^{\mathrm{L}} \end{cases} \end{aligned} \tag{47}$$

or, equivalently

$$\begin{aligned} &\tau_{\tilde{v}}^{k\,\circ}(v_1, \ldots, v_K, h, t_{\tilde{v}}) = \\ &\quad = \min\{s_k^{\mathrm{U}}, \max\{\lambda_{v_1,\ldots,v_K,h}^{k} - t_{\tilde{v}}, s_k^{\mathrm{L}}\}\} \end{aligned} \tag{48}$$

being $\tau_{\tilde{v}}^{k\,\circ}(v_1, \ldots, v_K, h, t_{\tilde{v}})$ the optimal duration of the stopover time, provided that T_k is selected as the next city to be visited after T_h.

3) Determine, for each (K+1)-tuple $(v_1, \ldots, v_K, h)$, such that $v_k \in \{0, 1\}$, $k = 1, \ldots, K$, $\tilde{v} < K$, and $h \in \{1, \ldots, K\}$, the conditioned costs-to-go and the optimal costs-to-go through the backward recursive relations

$$\begin{aligned} &\mathcal{J}_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}} = 1) = \\ &\qquad = \alpha_k \max\{t_{\tilde{v}} + \xi_{h,k} + \tau_{\tilde{v}} - d_k, 0\} + \\ &\qquad + \hat{\beta}(s_k^{\mathrm{U}} - \tau_{\tilde{v}}) + \\ &\qquad + \mathcal{J}_{v_1,\ldots,v_k+1,\ldots,v_K,k}^{\circ}(t_{\tilde{v}} + \xi_{h,k} + \tau_{\tilde{v}}) \end{aligned} \tag{49}$$

for each $k \in \mathcal{A}_{v_1,\ldots,v_K,h}$,

$$\begin{aligned} &\mathcal{J}_{v_1,\ldots,v_K,h}^{\circ}(t_{\tilde{v}}) = \\ &\quad = \min_{k \in \mathcal{A}_{v_1,\ldots,v_K,h}} \{\mathcal{J}_{v_1,\ldots,v_K,h}(t_{\tilde{v}} \mid \delta_{k,\tilde{v}} = 1)\} \end{aligned} \tag{50}$$

with initial condition $\mathcal{J}^{\circ}_{1,\dots,1,1}(t_K) = 0$.

4) Then, the optimal control strategies, for each (K+1)-tuple $(v_1, \dots, v_K, h)$, such that $v_k \in \{0, 1\}$, $k = 1, \dots, K$, $\tilde{v} < K$, and $h \in \{1, \dots, K\}$, are obtained as
   - if $\tilde{v} \le K - 3$:

$$\delta^{\circ}_{k,\tilde{v}}(v_1, \dots, v_K, h, t_{\tilde{v}}) = \\ = \begin{cases} 1 & \text{if } k = \operatorname{argmin}_p \{ \\ & \quad \mathcal{J}_{v_1,\dots,v_K,h}(t_{\tilde{v}} \mid \delta_{p,\tilde{v}} = 1), \\ & \quad p \in \mathcal{A}_{v_1,\dots,v_K,h} \} \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

   - if $\tilde{v} = K - 2$:

$$\delta^{\circ}_{k,\tilde{v}}(v_1, \dots, v_K, h, t_{\tilde{v}}) = \\ = \begin{cases} 1 & \text{if } k \in \{2, \dots, K\} \text{ and } v_k = 0 \\ 0 & \text{otherwise} \end{cases} \tag{52}$$

   - if $\tilde{v} = K - 1$:

$$\delta^{\circ}_{k,\tilde{v}}(v_1, \dots, v_K, h, t_{\tilde{v}}) = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

   and

$$\tau^{\circ}_{\tilde{v}}(v_1, \dots, v_K, h, t_{\tilde{v}}) = \\ = \sum_{k \in \mathcal{A}_{v_1,\dots,v_K,h}} \Big(\delta^{\circ}_{k,\tilde{v}}(v_1, \dots, v_K, h, t_{\tilde{v}}) \cdot \\ \cdot \tau^{k\,\circ}_{\tilde{v}}(v_1, \dots v_K, h, t_{\tilde{v}}) \Big) \tag{54}$$

Besides, the optimal cost-to-go $\mathcal{J}^{\circ}_{v_1,\dots,v_K,h}(t_{\tilde{v}})$ is a $\mathcal{PWL}(M, \hat{\beta})$-function, for some M, for any set $(v_1, \dots, v_K, h)$ such that $v_k \in \{0, 1\}$, $k = 1, \dots, K$, $\tilde{v} < K$, and $h \in \{1, \dots, K\}$. $\square$

VI. CONCLUSIONS

Two TSP models have been presented in this paper, and two constructive procedures for the determination of optimal control strategies have been proposed. In this way, it is possible to take in real-time optimal decisions, even in case of a deviation from the nominal system behaviour. Obviously, it does not mean that the proposed solution procedures overcome the difficulty of dealing with a NP-hard problem. In particular, the two procedures, which are based on the application of dynamic programming, require the determination of the discrete state space of the considered system, whose size has a non-polynomial dependence from the number of cities to be visited.

REFERENCES

[1] D.L. Applegate, R.E. Bixby, V. Chvátal, and W.J. Cook, *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, 2006.

[2] G. Gutin and A.P. Punnen (Eds.), *The Traveling Salesman Problem and Its Variations*. Combinatorial Optimization Series, vol. 12, Springer, 2007.

[3] G. Gutin and A.P. Punnen, "The traveling salesman problem", *Discrete Optimization* (Special Issue), vol. 3, no. 1, 2006.

[4] N. Christofides, "Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem", Internal Research Report, Carnegie-Mellon University, Pittsburgh, PA (USA), 1976.

[5] E. Balas, P. Toth, "Branch and Bound Methods for the Traveling Salesman Problem", in G. Lawler, J.K. Lenstra, A.H.G. Rinnoy Kan, and D. Shmoys (Eds.), *The Traveling Salesman Problem*, John Wiley, 1985.

[6] M. Dorigo and L.M. Gambardella, "Ant colonies for the travelling salesman problem", *BioSystems*, vol. 43, 1997, pp. 73-81.

[7] M. Aicardi, D. Giglio, and R. Minciardi, "Optimal strategies for real-time determination of the next job's class in a single machine with setup times and controllable processing times", *Proceedings of the European Control Conference 2007*, Kos (Greece), 2007, pp. 3963-3968.

[8] M. Aicardi, D. Giglio, and R. Minciardi, "Optimal Strategies for Multiclass Job Scheduling on a Single machine With Controllable Processing Times", *IEEE Transactions on Automatic Control*, vol. 53, no. 2, 2008, pp. 479-495.

[9] R.G. Vickson, "Choosing the job sequence and processing times to minimize total processing plus flow cost on a single machine", *Operations Research*, vol. 28, no. 5, 1980, pp. 1155-1167.

[10] N.G. Hall, S.P. Sethi, and C. Sriskandarajah, "On the complexity of generalized due date scheduling problems", *European Journal of Operation Research*, vol. 51, no. 1, 1991, pp.100-109.