IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received 20 May 2024, accepted 29 July 2024, date of publication 12 August 2024, date of current version 23 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3441710

RESEARCH ARTICLE

Including the Temporal Dimension in the Generation of Personalized Itinerary Recommendations

FEDERICA CENA, LUCA CONSOLE, MARTA MICHELI, AND FABIANA VERNERO
Department of Computer Science, University of Turin, 10149 Turin, Italy
Corresponding author: Marta Micheli (marta.micheli@unito.it)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Bioethical Committee of the University of Turin under Protocol No. 0631962.

ABSTRACT Tourists spend a lot of effort in planning itineraries when organizing a trip. This is a complex activity that involves selecting the places to visit and dealing with a number of temporal issues to generate a schedule for the visit. In the paper we propose an architecture for a recommender that suggests personalized tourist itineraries and a personalized time schedule. The approach takes into account (i) user preferences for the places to be included in the itinerary and (ii) several temporal dimensions concerning both temporal information and constraints (e.g., opening hours, time for visiting each place, time to move among places) and time-related user preferences (e.g., number of days of the visit, preferring a dense schedule vs having a lot of free time, the variety of the types of attractions during a day of visit). The approach is based on a combination of genetic algorithms and temporal reasoning. It focuses on generating temporally annotated itineraries starting from a ranked list of places to be included. Thus, our solution is designed as a module that can be coupled with any system recommending places to visit. The design started from a user study we carried out to analyze the temporal dimensions to take into account, and to find relationships between such dimensions and users' personality traits, and led to the development and evaluation of a prototypical implementation that generates personalized itineraries for the city of Turin.

INDEX TERMS Itinerary recommendation, recommender system, temporal information.

I. INTRODUCTION

Tourists spend a lot of effort in planning itineraries when organizing a trip, starting from the selection of the places they would like to visit and then scheduling the visits based on a variety of temporal constraints and preferences such as the opening times of the places, the visiting time for each attraction, the time needed to move among places, the time to rest and have meals.

Nowadays, there is a plethora of information available on websites, travel guides and magazines concerning tourist attractions. This massive volume of information makes it a challenging and cumbersome task for tourists to process all potential options. Thus, developing solutions that enhance the use of such information and help users make decisions based on their personal interests and constraints has become essential [1].

Recommender Systems *"produce individualized recommendations as output or have the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options"* [2]. They have been used to support people in decision-making processes in different domains (e-commerce, movies, books, tourism). In the context of tourism, to provide effective support, a recommender system has to manage a large amount of information about users, places and context, as well as to take into account all the different aspects that can impact the decision-making processes.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Shorif Uddin.

More specifically, tourist recommenders should be able to offer solutions for the *itinerary planning problem* [3], which refers to the planning tasks for tourists interested in visiting multiple points of interest (PoI in the rest of the paper) with the objective of maximizing tourist profit. This task is especially complex and challenging due to: (i) the need to identify a set of attractions that are also aligned with the traveller's interests; (ii) the need to organize these attractions as an itinerary compliant with the constraints to move among places and limited time for touring; and (iii) the need to plan for travelling, visiting and queuing times at the attractions, where queuing times are dependent on the time of the attraction visit. In particular, considering crowded hours is crucial: neglecting to consider possible queuing times can create a frustrating experience for travellers as they spend a long time queuing instead of enjoying the attractions, and possibly miss attraction visits in their itineraries due to these queuing times exceeding their available touring time [4].

Hence, time is a crucial aspect to be managed in recommenders for tourism, especially in the case of itinerary recommendations [5], [6], [7]. In this case, to provide an optimal user experience, it is crucial to consider also the user preferences about all the temporal-related features a trip can have. These features include travelling time, preference for having a dense schedule or a lot of free time, preference for visiting many places quickly or, at the opposite, visiting fewer places in depth, preference for alternating the type of places (e.g., museums, parks, …), etc. However, time is usually considered only as a *constraint* rather than an *object of user preference*, thus not considering that the peculiar needs of a person that can be very different from those of another one. To our knowledge, very few works consider user preferences on specific temporal dimensions: [8], [9] take into account user preferences for the travelled distance, and [10], [11] consider individual preferences for visit duration. User preferences for all the temporal-related features of a trip have not been considered in itinerary recommender systems so far.

The aim of this paper is to fill this gap by proposing an innovative approach and architecture to tackle the problem of recommending personalized itineraries, taking into account user preferences for all the temporal dimensions mentioned above.

In particular, in order to cope with temporal information and all the user constraints and preferences together, we decided to design an approach that integrates genetic algorithms and temporal reasoning: the former is in charge of building candidate personalized itineraries which consider the user preferences on temporal features, while the latter is in charge of checking the temporal constraints in the itinerary producing a detailed time schedule.

Genetic algorithms (GAs) are very suitable for the problem we are facing, which can be seen as a constrained optimization one. Recently, the adoption of evolutionary algorithms has been considered in a number of works to deal with the problem of generating and recommending touristic itineraries [3], [12], [13], [14], [15]. A peculiarity of our approach is that the temporal preference dimension will play an essential role in the evaluation function of the genetic algorithm and constraint satisfaction will be used to prune itineraries that violate them and thus are not temporally consistent.

In particular, we chose to adopt an approach based on GA for two main reasons. First, it can deal in a simple, natural and integrated way with the combinatorial nature of the problem we are facing. Moreover, by varying the fitness measure, it allows to experiment different ways of combining temporally related user preferences. Second, the approach is any-time, and thus it can generate some sort of temporal recommendation even in a very short time and with limited computational resources, being suitable for assembling itineraries on the fly.

In the paper, we describe the architecture of the itinerary recommender, its application in the design of a prototype, and the evaluation of our approach. The main contributions of this work can be summarized as follows:

- We propose a modular architecture that can be coupled with any recommender that produces a ranked list of places that are best suited for a given user. The advantages of this choice as regards both the design and evaluation of the approach will be discussed in detail.
- We take into account the temporal dimension in a thorough manner, considering all the facets where time can play a relevant role, both from an objective (time related to places and routes, such as distances, opening hours) and subjective point of view (user preferences for a specific facet, i.e. preferences for specific opening hours).
- Temporal information and reasoning is used in two ways: on the one hand to check the temporal consistency of itineraries and thus produce realistic time schedules; on the other hand, to evaluate the suitability of an itinerary for a user to provide personalized recommendations.
- We performed a thorough user study to assess the advantages produced by personalized temporal itinerary recommendations.

A preliminary description of the work has been provided in [16]; that paper, however, included a very preliminary test with users and did not include the experiment discussed in this paper. Moreover, the algorithm itself has been revised and refined given the results of the preliminary evaluation of the prototype. Section VI includes more detailed specifications regarding this modification.

The paper is structured as follows. In the next Section, we provide an overview of the relevant state-of-the-art work. Section III includes an introduction to the proposed recommender system, whose information requirements are detailed in Section IV. We illustrate the design and implementation of our solution in Sections V and VI. Then, Section VII introduces the prototype we developed and

Section VIII provides information on the recommender system evaluation. Finally, in the last part of this article, we discuss the limitations of the work and possible directions of improvement.

II. RELATED WORK

A. ITINERARY RECOMMENDER SYSTEMS

Several approaches in the literature recommend not only single PoI but also a complete combination of a set of PoIs. This can be done taking into account several features of the path, such as efficiency (length and speed) [17], [18], but also pleasantness [19], accessibility [20] and safety [21].

Many works on itinerary recommendation have the main objective of recommending an itinerary that maximizes a global profit/reward and can be completed within a specific budget [22], [23], [24].

In recent years, many works have incorporated *user interests and/or specific preferences* to personalize the construction of itinerary [1], [3], [4], [25], [26], [27], [28], [29], [30], [31]. Other works aim to recommend itineraries that consider personalized requirements such as a specific attraction visit sequence [32], mandatory attraction categories [10] or group interest satisfaction [33], [34].

Managing time is very important in itinerary recommenders [5], but it is often seen as a constraint rather than as an *object of user preference*, such as where they want to stay or what time they want to go.

Of course, determining the proper visiting time of each place and the proper transit time from one place to another is fundamental for defining route goodness functions [35]. Techniques based on signal processing are proposed for including time dimension in context-aware recommendation tasks [36], [37], [38]. Other works, such as [39] and [40], aim to first identify a set of interesting and popular attractions, then construct an itinerary that comprises these attractions using a variant of the Travelling Salesman Problem [41].

However, very few works consider user preferences on temporal dimensions. We can cite for example [8] and [9], that construct personalized itineraries optimizing attraction popularity and user interests relative to the travelled distance, and [10], [11] that recommend popular attractions tailoring the visit duration based on user interest.

Apart from *how* time can be managed by the recommendation algorithm, time itself can be considered as a multidimensional aspect, with different facets as presented in [42].

In the following, we describe which temporal dimensions have been considered by some of the most relevant itinerary recommenders in the literature.

Di Bitonto et al. [31] propose a method for generating tourist itineraries in knowledge-based recommender systems. The method is based on a theoretical model that defines space-time relations among items of intangible cultural heritage (called events) and on transitive closure computation of the relations, that is able to construct chains of events. The output is a sequence of attractions or spots to be visited, filtered according to the tourist's constraints (day of visit, cost, and so on) specified in the request.

Yoon et al. [43] explicitly model both the available time of the user and the staying time for each PoI included in the itinerary.

Refanidis et al. [25] present an intelligent web-based system aiming at making recommendations based on several place-related features (like location, cost, availability, duration range, visiting time) and user-selected criteria, such as visit duration and timing, geographical areas of interest and visit profiling.

Lim et al. [10] propose an algorithm for recommending personalized tours using PoI popularity and user interest preferences, which are automatically derived from real-life travel sequences based on geotagged photos. They consider user trip constraints such as time limits. In our work, we also reflect levels of user interest based on visit durations.

Fogli et al. [28] present a recommendation engine that considers the user profile, the context of use, and the features of the PoIs extracted from linked open data (LOD) sources, and, in relation to time, total available time, travel time, visiting time, opening time and distance between PoIs.

Cai et al. [29] propose an itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos by discovering sequential PoIs with temporal information (travel time, opening time) from other users' visiting sequences and preferences.

Taylor et al. [1] propose and formulate the TourMustSee problem (based on a variant of the Orienteering problem), which incorporates a set of must-see PoIs into travel itineraries, along with considerations of a starting/ending PoIs and travel times between PoIs and visit durations at PoIs.

Zou et al. [44] propose a recommendation system that, according to the requirements specified by its users, is capable of generating a few travel itineraries of distinct features (e.g. the most appealing itinerary, the shortest itinerary or the itinerary with the highest performance/price ratio).

Tenemaza et al. [3] consider several aspects in the recommendations: the context of a tourist destination visited, lack of updated information about PoIs, transport information, weather forecast, available time, travel time and visiting time. They present a mobile recommender system based on Tourist Trip Design Problem (TTDP) (Time Depending Orienteering Problem with Time Windows), which analyzes in real time the constraints related to users and PoIs, and implements a genetic algorithm.

Chen et al. [26] propose a framework to infer the user interests and recommend personalized itinerary consisting of PoIs, visit durations, visit sequence and total available time.

Ji et al. [45] address the problem of considering the attractions' spacial heterogeneity when creating personalized trips. In their heuristic-based approach, an improved artificial bee colony algorithm and a differential evolution algorithm are adopted to generate itineraries, optimized according to the visiting time duration.

Erbil et al. [46] split the problem of creating multi-day trips into two parts: the list of PoIs to be visited each day is first created using agglomerative clustering, and then the lists are ordered using a greeding approximation algorithm avoiding PoIs from the same category to be placed next to each other. In addition to the user preference for PoI categories, this solution also considers the traveller's pace and visiting effort.

Halder et al. [47] propose to use the Monte Carlo Tree Search algorithm to design itineraries considering the user interest, attraction popularity, visiting time and queuing at the PoIs. To avoid non-optimal itineraries and reduce the solution space, they also introduce a pruning technique.

Our solution differs from these existing works since it produces personalised tourist itineraries taking into account a wide range of time-related constraints (total available time, free time, lunch breaks, transfer times, PoIs' opening hours, duration of the visits, busy hour avoidance) as well as other user preferences (mandatory PoIs, variety of PoIs, PoIs' ranking, quantity of PoIs). According to our analysis, no other work takes all these time-related features into account altogether.

B. GENETIC-BASED RECOMMENDER SYSTEMS

Recently, the adoption of genetic algorithms has been considered in a number of works to deal with the problem of generating and recommending tourist itineraries. The solutions proposed so far differ from each other in terms of optimization objectives, parameter settings of the algorithm used, and adoption of modules other than the genetic algorithm.

One of the first works in which a genetic approach has been used to solve the itinerary planning problem is that of Chen et al. [12]. In their system, a list of recommended POIs is produced through the item-based collaborative filtering method, and then a genetic algorithm is adopted to find an ordered set of tourism places given the total time available and travel budget restrictions.

Karbowska-Chilinska and Zabielski [48] introduce time windows in a genetic algorithm for route generation, ensuring that visit schedules take into account the POIs' opening and closing times. Their approach maximizes the sum of the profits associated to each location in the route.

Changdar et al. [49] adapt a generative algorithm to address the Travelling Salesman Problem, which involves finding the shortest path to visit a set of cities exactly once, starting and ending the tour at the same place. Fuzzy total travel time and cost are considered in their work. In [50], the same authors address again this problem, extending it to multiple salesmen by combining a genetic algorithm with the ant colony optimization technique.

Wibowo and Handayani [14] include restaurant selection in a genetic-based travel itinerary recommender. While their algorithm overlooks user preferences, such as the desired duration of the itineraries, it considers POIs' opening hours, scheduling selected restaurants at lunch and dinner time.

Yochum et al. [15] personalize tourist itineraries using an adaptive GA based on the quantity of POIs present in the solution, their popularity, overall rating, cost, and type (mandatory or not). Similar to our work, they consider user preference for multiple factors. However, POIs' opening hours and other time-related constraints (e.g., lunch breaks) that could heavily impact the user's ability to adhere to the itinerary are not taken into consideration.

Qomariyah and Kazakov [51] use a GA, but only to identify the set of destinations that best fit the travel duration constraint and the user budget. A separate layer based on Google Optimization Tools is employed to connect the selected destinations into a single itinerary.

Some works accompany genetic algorithms with other techniques. For example, Zheng et al. [13] combine a GA with a different evolution algorithm to generate single-day itineraries. In their system, tourist aesthetic fatigue, associated to the POI visiting duration, and the sightseeing value of attractions are both considered.

Tenemaza et al. [3] separate the itinerary recommendation problem in two parts. In their work, a k-means algorithm is first used to clusterize POIs depending on the number of available visiting days, and then a genetic algorithm is adopted to minimize the difference between the visiting time and the total available time. The evaluation of feasible itineraries also takes into account the violation of POIs' opening hours, modeled as penalties.

Ghobadi et al. [52] also explore the concept of integrating genetic algorithms with other techniques. They adapt a hybrid algorithm, previously introduced by [53], to tackle the multi-day trip planning problem. Their approach combines a genetic algorithm and a variable neighborhood descent algorithm to generate itineraries, focusing on preserving diversity among the recommended POIs.

With respect to related works using genetic-based approaches, we generate and evaluate itineraries while taking into account both user preferences for a wide number of contextual aspects and the imposition of time-related constraints, such as the opening hours of POIs or the visit durations, through a dedicated validation module. In our algorithm the maximization or minimization of specific time-related factors is not standardized. Instead, it adapts based on user preferences, exerting differing levels of influence on the resulting itinerary. Furthermore, when recommending multi-day itineraries, we propose to consider balance across the days, prioritizing itineraries that offer a satisfying and attractive experience each day, without favoring any day over the others.

III. ASSEMBLING A TEMPORAL ITINERARY

As we noted in the introduction, we propose to decouple and serialize two aspects of the problem of generating personalized itineraries:

1) *Selecting items to be recommended.* Generating a-temporal recommendations, i.e., producing a ranked

list of suggested PoIs that are most suitable for a given user.

2) *Itinerary construction.* Building a temporal itinerary, starting from the results of the previous phase.

We claim that this approach can have significant advantages.

First of all, it allows us to decouple two complex problems that indeed rely on different information (knowledge) sources and that in this way can be dealt with independently.

A second advantage is that in this way our itinerary construction module can be coupled with any recommender, based on any technology. The module, in fact, will only assume that a ranked list of items is produced as an effect of the first step, regardless of the fact that this list is generated using collaborative or content-based filtering, knowledge-based recommendations or other approaches.

A third advantage is that we can evaluate the approach that generates the itineraries *per se*, independently of the selection of the items that are of interest for the user. In our evaluation, in particular, we will assume to have this list directly from users, as some sort of ''ground truth'' which we do not expect to play a confounding role in the evaluation of recommended itineraries. Thus we can assess the ''actual'' advantages of the approach.

Finally, the decoupling will allow us to exploit a number of results and approaches from the literature on temporal reasoning and temporal constraint solving.

IV. THE TEMPORAL KNOWLEDGE BASE

In this section we discuss the temporal dimensions that we take into account in the generation of itineraries and we analyze the structure of the temporal knowledge base. This involves two different aspects, (i) temporal information about the PoIs that will constitute the temporal knowledge base and (ii) temporal information and preferences concerning the user that will constitute the temporal user model.

A. TEMPORAL KNOWLEDGE ABOUT THE POINTS OF INTERESTS

In this section, we discuss the information about PoIs which is needed for generating the itineraries, analyzing how these pieces of information can be made available and represented.

- Spatial coordinates of the PoI. Although this piece of information is not strictly temporal, it is necessary to compute the temporal distances between the PoIs. Coordinates and temporal distances are easily available using any of the map applications on the Internet. They can provide different types of temporal distances (walking, by bike, by car, by public transport). The temporal distance between two PoIs can be represented either as a precise numeric value t or as a pair $[t_{min}, t_{max}]$ representing the minimum and maximum distance. We can also consider multiple temporal distances in case the transfer between the point is performed on foot, by car or by public transport.
- Opening hours for each PoI, for each day of the week. Also, this piece of information is commonly available on the Internet. In case it is not available, assumptions can be made based on the type of PoI and the location (e.g., typically museums in Italy are open 9-18 every day but Mondays). Opening and closing times are then expressed as points t_{open} and t_{close} on the timeline.
- Crowding moment estimation across time and days. This piece of information is not strictly necessary, as we shall see, but can be used to optimize the itinerary, especially in case a user prefers to avoid visiting places when there are too many people. Crowding estimations are available on the Web for many PoIs. We consider three qualitative values $\{low, medium, high\}$ for crowding and then for each day of the week we divide the opening time into intervals for which crowding is low, medium or high.
- The time for visiting the PoI, expressed either as an average duration or as an interval $[t_{min}, t_{max}]$ distinguishing between minimum time for a quick view to a maximum time for an extensive visit. Alternatively, the average time visitors spend could be used, if available. Also in this case the information is available for some PoIs or can be estimated.
- Availability of services that may be important at some times of the day (e.g., cafes, restaurants, ...) and their opening times.
- Moreover, each PoI is annotated with the type, or category, of attraction (e.g., museum, park, historic building, ...).

B. TEMPORAL INFORMATION ABOUT THE USER AND USER TEMPORAL PREFERENCES

Let us analyze now the pieces of temporal information that we expect from the user for whom the recommendation (itinerary) has to be generated. The pieces of information below will constitute the temporal user model that will be exploited by the algorithm for generating and ranking itineraries and thus for providing personalized itinerary recommendations.

- Time available for visiting the location for which recommendations (itinerary) have to be generated. We expect that the user provides information about the day and time of arrival t_{start} and the day and time of departure t_{end} from which we can determine the duration of the visit (in hours, days, ...).
- User desire to carve out some free time during the trip, if available. Indeed, a tourist may want to spend some time in activities other than visiting PoIs (e.g., shopping, relaxing). In this case, some free time will be introduced in the recommended itinerary.
- User preference for the amount of time to be devoted to the visit of each PoI. Some people may be interested in visiting as many places as possible, devoting little time to each visit, while others may prefer to visit fewer PoIs but devote time to a detailed visit of each of them.

This preference could also distinguish between different types of PoIs and, for example, a user could express the preference for a thorough visit of museums and a quick visit to natural sites such as parks.

- User interest in visiting heterogeneous PoIs. If this information is provided by the individual, the recommender system will include different types of PoIs in the itinerary (e.g., natural parks, historical museums, religious places, etc.).
- User preferences about the minimization of transfer times. Some people may prefer to spend as little time as possible travelling from one place to another, while for others it may not be so important, preferring, for example, to travel long distances if a place is particularly interesting.
- User preferences about crowding, if available. In particular, it could be interesting and important to have this piece of information for users who suffer from staying in crowded places. If not available, the planner will prefer less crowded times of visit over crowded ones and it will anyway warn the user in case an itinerary includes visiting crowded PoIs.

V. CHARACTERIZING TEMPORAL ITINERARIES AND THEIR GENERATION

In this section we discuss the algorithm that generates a ranked list of itineraries that are most suitable for a given user. The approach relies on a genetic algorithm which exploits constraint satisfaction in the evaluation process, as we will clarify in the following section. Constraints, in particular, are expressed as bounds on differences [54] which is thus the formalism we adopted to represent all the temporal information associated with an itinerary.

The input of the process is:

1) A ranked list $REC\_ITEM = \{\langle Item_1, V_1\rangle, \langle Item_2, V_2\rangle, \ldots, \langle Item_n, V_n\rangle\}$ of the PoIs as suggested by the recommendation component, where:
   - Each $Item_i$ is a PoI;
   - $V_i \in [0, 1]$ is an evaluation of the suitability of $Item_i$ for the current user (anyhow it is produced).

   The user can also specify a list of *must see* PoIs that should be included in all the itineraries.
2) The starting t_{start} and ending point t_{end} (date and time) of the itinerary to be planned.
3) Temporal information about the PoIs, as discussed in Section IV-B.
4) Temporal user information and possibly preferences, as discussed in Section IV-B.

The genetic algorithm maintains a population formed by a (ranked) set of valid itineraries:

Definition 1: A Population P is a ranked set:
$P = \{\langle It_1, f(It_1)\rangle, \langle It_2, f(It_2)\rangle, \ldots \langle It_n, f(It_n)\rangle\}$, where:

- each It_i is a valid itinerary
- $f(It_i)$ is its numeric evaluation

where an itinerary It is defined as follows:

Definition 2: An itinerary It a sequence of items:
$It = \langle\langle Item_1, t_{s1}, t_{f1}\rangle, \langle Item_2, t_{s2}, t_{f2}\rangle, \ldots \langle Item_m, t_{sm}, t_{fm}\rangle\rangle$ where

- each $Item_x$ is a Place of Interest (PoI)
- t_{sx} and t_{fx} are the starting and ending time points of the visit of $Item_x$.

Only the itineraries that are consistent with the temporal knowledge base are maintained. In particular,

Definition 3: An itinerary
$It = \langle\langle Item_1, t_{s1}, t_{f1}\rangle, \langle Item_2, t_{s2}, t_{f2}\rangle, \ldots \langle Item_m, t_{sm}, t_{fm}\rangle\rangle$
is **valid** if and only if it is temporally consistent, i.e., the set of constraints at the following items are consistent:

- The visit of each item must be during the opening times of the item itself
  $t_{open}(Item_x) \leq t_{sx}$
  $t_{fx} \leq t_{close}(Item_x)$
- visit time for each $Item_x$ is within the minimum and maximum time specified by the temporal knowledge base
  $min\_visit\_time(Item_x) \leq t_{fx} - t_{sx} \leq max\_visit\_time(Item_x)$
- time to move between consecutive items $Item_i$ and $Item_j$ is greater than the transfer time
  $t_{sj} - t_{fi} \geq time\_distance(Item_i, Item_j)$
- the starting time of the visit of the first item is greater than the starting point of the itinerary
  $t_{s1} \geq t_{start}$
- the end time of the visit of the last item is less than the ending point of the itinerary
  $t_{fm} \leq t_{end}$

This means that the time interval $[t_{sx}, t_{fx}]$ allocated to each $item_x$ must be consistent with its opening times and with the time needed to visit it, and that for each pair of consecutive items the time to move from one to the following must be consistent with transfer times. These constraints can be expressed as bounds on differences on the variables t_{si}, t_{fi}; thus validity can be checked with any temporal solver working on bounds on differences. In this way the solver can also produce consistency intervals for each variable and thus a schedule for the visit t_{si}, t_{fi} of each item. The approach in [55] allows us to perform this check in a very efficient way.

For each valid itinerary It the evaluation f_{It} is computed taking into account the two following aspects:

- the evaluation of the items included in the itinerary, as it is produced by the recommender and thus tailored to the individual user;
- the user temporal preferences.

This is part of the genetic algorithm and will be discussed in detail in the following section.

VI. A GENETIC ALGORITHM FOR ITINERARY CONSTRUCTION

Genetic algorithms are a subclass of evolutionary algorithms, inspired by natural evolution theories and biological genetic.

The idea of evolution in GAs is actualized through the adoption of generations that follow each other until a desired solution is found. Each generation - or population - is composed by both individuals (solutions) of the previous generation and offsprings. Indeed, members of a generation can reproduce to obtain a new individual which is the result of two individuals merged together [56].

Our genetic algorithm:

- Starts with an initial population of candidate valid itineraries generated in a random way with an element from *REC_ITEM*. Each itinerary includes a number of days, based on the information provided by the user on the duration of the visit. The itineraries are ranked.
- Iterates making the population evolve across generations. A new population is generated by computing descendants of the current one (crossover) and introducing mutations. The itineraries that are not valid are removed.
- Stops when the evaluation of the best-ranked itineraries stabilizes (or after a maximum preset number of generations).

The set of complete valid itineraries that are produced can be presented to the user as an ordered set of suggestions.

Depending on the available information on the user, different factors are taken into account by the algorithm, affecting the evaluation of individuals. Hence, our algorithm's output is an itinerary which is personalized according to user preferences. The adoption of a genetic algorithm makes our recommender system flexible, since it allows to select, combine together and balance multiple user preferences instead of optimizing one factor a time.

In order to determine what characteristics the desired solution should have, each factor is assigned a score *Upref* in the range [0, 1] that represents the user preference for that specific feature. If no information is available for a certain factor, its value is automatically set to 0 unless specified differently, implying that the factor will not influence the algorithm output.

The user model concerning temporal preferences is thus characterized by the five dimensions mentioned above, each one with an associated value $Upref_i \in [0, 1]$: preferences for (i) dense vs sparse itineraries (number of PoIs in the itinerary, $Upref_1$), (ii) having free time ($Upref_2$), (iii) avoiding visiting PoIs during crowded times ($Upref_3$), (iv) heterogeneous itineraries ($Upref_4$) and (v) minimizing transfer times ($Upref_5$). In the prototype we directly elicited such values from the user. Alternatively, it is possible to derive them from related information the recommender might already have. For example, we carried out a user study that showed how preferences for temporal dimensions can be determined based on users' personality traits [42].

Let us now summarize how our system works, introducing the various steps that compose the genetic algorithm and its interaction with constraint solving.

First, the system receives information on the user's date and time of arrival and departure. For each visiting day, some unavailable intervals are generated, including those for nights, lunch breaks and free time, if the user prefers to have some. In the latter case, the starting time and the duration of free time intervals are chosen randomly, in order to ensure variety between individuals in the initial population.

A. INITIALIZATION

The first generation is produced randomly to ensure sufficient diversity and, thus, to avoid premature convergence to a local optimum. Whenever a new individual (itinerary) has to be produced, the algorithm picks up PoIs from the list *REC_ITEM* taking into account their evaluation, i.e., PoIs that have a better evaluation have a higher probability of being selected: starting from the first one, the algorithm tries to place each attraction on the timeline in the first available interval. If the user explicitly chooses some PoIs that s/he wants to visit (*must see* PoIs), these are added to the top of the list, so that their probability of being inserted in the itinerary is higher. When no more attractions can be appended to the itinerary, the itinerary is considered complete.

At this point (i) temporal constraint checking is performed and (ii) if an itinerary is valid, it enters the first generation and its fitness is evaluated.

B. FITNESS EVALUATION

In order to quantify their goodness, each individual *It* in a population is assigned a fitness value:

$f(It) = f_{PoI}(It) + f_{Tpref}(It)$

which is the sum of two parts referring respectively to the PoIs included in *It* ($f_{PoI}(It)$) and to the evaluation of temporal preferences on *It* ($f_{Tpref}(It)$).

- The evaluation $f_{PoI}(It)$ of the PoIs in the itinerary according to the ranking in the list *REC_ITEM*. In this way, an itinerary containing PoIs that have a higher ranking receives a better evaluation. In particular, we consider two factors: (i) the average of the evaluation of the PoIs in the itinerary and (ii) the percentage of *must see* PoIs in the itinerary:

$$f_{PoI}(It) = average(V_{item_1}, V_{Item_2}, \ldots V_{Item_m}) * \frac{number\ of\ must\ see\ PoIs\ in\ It}{number\ of\ must\ see\ PoIs} \quad (1)$$

where V_{item_i} is the evaluation of the PoI $Item_i$.
Initially, the evaluation of the POIs in the itinerary did not include *must see* POIs, which were integrated after a preliminary evaluation of the algorithm presented in [16]. Indeed, feedback from several participants in the user tests, provided through open-ended responses, highlighted that various POIs they had designated as highly interesting to visit were not included in the recommended itineraries. Consequently, we placed greater emphasis on their preferences for favourite POIs by introducing the selection ratio of *must see* POIs in the evaluation function $f_{PoI}(It)$.

We also considered the possibility of pruning the itineraries whose evaluation $f_{PoI}(It)$ is below a threshold.

- The evaluation $f_{Tpref}(It)$ of the itinerary according to the user temporal preferences. Each type of preference is weighed depending on how important it is for the user. Each type of preference is considered separately and thus $f_{Tpref}(It) = \Sigma_{i=1}^{5}(f_{Tpref_i}(It) * Upref_i)$ where each $f_{Tpref_i}(It)$ corresponds to a temporal preference and is computed (using a number of heuristic evaluation criteria discussed below) as in the following five sub-items, and $Upref_i$ is the corresponding user model value for that temporal preference, represented on a scale [0, 1].
  - The itineraries are ranked according to the number of PoIs they contain; this ranking contributes positively to the evaluation of an itinerary if the user expressed a preference for maximizing the number of PoIs during the visit. $f_{Tpref_1}(It) = \frac{number\ of\ PoIs\ in\ It}{total\ number\ of\ PoIs}$
  - Similarly, itineraries are ranked as regards the amount of free time and also in this case the ranking can contribute positively or negatively based on the user preference. $f_{Tpref_2}(It) = \frac{sum\ of\ duration\ of\ free\ time\ intervals\ in\ It}{duration\ of\ It}$
  - The allocation of PoI visits to busy hours can negatively impact the evaluation if the user prefers to avoid crowding. $f_{Tpref_3}(It) = \frac{number\ of\ PoIs\ visted\ during\ crowded\ times\ in\ It}{number\ ofPoIs\ in\ It}$
  - The heterogeneity of PoIs contributes positively if the user prefers this type of itinerary, negatively otherwise. $f_{Tpref_4}(It) = number\ of\ different\ types\ of\ PoIs\ in\ It$
  - Furthermore, transfer time is taken into account and if the user prefers its minimization, the itineraries are assigned different ranking values based on the sum of transfer times. $f_{Tpref_5}(It) = \frac{sum\ of\ transfer\ times\ in\ It}{duration\ of\ It}$ Notice that this corresponds to a global analysis of transfer time during the whole itinerary and does not necessarily imply the minimization of each transfer.

  Finally, in case the visit includes multiple days, we avoid fitness imbalances between the different days, calculating the fitness of the individual f_{It} as follows: $f(It) = tf - \sqrt{\frac{\Sigma_{i=1}^{N}(f_i - \mu)^2}{N}}$ where tf is the initial total fitness of the whole individual, f_i are the fitness values of the single days of the itinerary, μ represents their average and N is the number of days.

C. TEMPORAL CONSTRAINT CHECKING

Each time a new event is added to an individual or an individual is modified, a validation module is used by the system to check the validity of the temporal constraints. An itinerary is temporally consistent if PoIs are open during the whole duration of the planned visit, the duration itself is within the minimum and maximum estimated visit time, and the time to transfer from a PoI to the next one is greater than the minimum transfer time. More precisely, a set of constraints (bounds on differences between time points) is created according to definition 3 based on the temporal knowledge base and checked. This can be done efficiently starting from the preprocessing of the constraints in the temporal knowledge base exploiting the results in [55].

If the individual is valid, the constraint propagation algorithm produces validity intervals for each time point in the itinerary (starting and ending point of each item in the itinerary). Thus when an itinerary is presented as a solution, it carries on detailed temporal information (expressed as a time interval on the timeline) on when the visit of each item is scheduled.

D. ITERATION: PRODUCING A NEW GENERATION

The algorithm iterates producing subsequent generations of the population of itineraries. Itineraries in the current generation are selected and combined (crossover). Mutations are introduced periodically. Let us analyze these steps in more detail.

1) SELECTION

After fitness values are computed for all the itineraries, the system is ready to identify what individuals will be used to generate the next population. A mating pool is produced using the Roulette-wheel Selection Method that allows to obtain a fitness proportionate selection of individuals. According to this method, the probability p of selection for each individual It of a population P results from the following formula:

$p_{It} = \frac{f_{It}}{\Sigma_{j=1}^{N} f_j}$ where N is the number of individuals in the existing population.

2) CROSSOVER

Two parents It_1 and It_2 are randomly selected from the mating pool and, according to the probability defined by the crossover rate, parents are copied in the next generation or are merged to generate an offspring It_3. This process is repeated until a new complete population is obtained. Temporal propagation and checking is performed on It_3, which enters the new population only if it is valid.

Crossover can be performed using various techniques, differing in the way each individual is split and mated. Different splits of It_1 and It_2 can be considered (e.g., split in two parts, multiple split of each day at half, multiple split at the end of each day, ...) and consequently different strategies for recomposing the parts in I_3.

Two types of crossover have been tested to determine what is the best for the proposed solution: a one-point crossover, whose crossover point is represented by a single event at the middle of the itinerary, and, for multi-day itineraries, also an N-point crossover in which a crossover point for each day is selected. In our study, better results have been observed using the one-point crossover.

3) MUTATION

Each new candidate itinerary can be subject to a mutation, with a probability dictated by a mutation rate value. When mutation if performed, a PoI is randomly selected from the individual and is replaced by another PoI that has not yet been included in the itinerary. Mutation covers an important role in the generation of near-optimal solutions because it is used to maintain genetic diversity in the population and thus to avoid local minima [57]. Mutated individuals must pass the temporal consistency check to enter the population.

After mutations are performed, fitness values are calculated for the itineraries in the new population and all the steps of the genetic algorithm (except for Initialization) are performed again until termination conditions are reached.

E. TERMINATION AND SOLUTIONS

The algorithm proceeds by producing subsequent generations. Two alternative strategies can be considered for terminating the iterative production of new generations: (i) deciding a-priori the number of iterations (generations) to be produced, possibly tuning this number after some experiments and (ii) iterating until the evaluation of the best ranked itineraries in the population does not change significantly from one iteration to the next one. In our prototype we selected the first strategy.

At this point, a solution is presented to the user in the form of a ranked list of the itineraries with best fitness.

F. TUNING

Alternative strategies can be adopted in various parts of the algorithm: the optimal values for the crossover rate and the mutation rate usually need to be determined by trial and error. The choice of these parameters is essential for the effectiveness of GAs [58]. Also the size of the initial population and the number of iterations are factors to be determined in advance. We performed this tuning exploiting the prototype that will be described in the next sections: several tests have been conducted to identify the parameter values that best suit our system's characteristics. After executing the GA for over 20000 times, the following optimal parameter setting was found (including also the type of crossover, as explained in Section VI-D):

- crossover rate: 0.8
- mutation rate: 0.6
- number of iterations: 20
- first population's size: 30
- crossover: one-point crossover for both single-day and multi-day itineraries.

We do not claim these settings are optimal in general, since further tests on other applications should be performed; however, they represent useful guidelines for the further development of our approach.

During the tuning process we also made a few experiments with the heuristics for evaluating the temporal suitability of itineraries and we chose those described in Section VI-B.

VII. PROTOTYPE

In order to evaluate our approach, we developed a prototype which provides personalized itinerary recommendations for the town of Turin through a web interface. The prototype has been developed with two goals in mind: (i) making experiments with alternative strategies and tuning the genetic algorithm (see the discussion above) and (ii) performing an overall evaluation of the approach (see Section VIII).

In order to populate the prototype with content, we developed a knowledge base with 30 PoIs in Turin, including a variety of attractions (museums, parks, churches, historic buildings, historic squares . . .).

We then collected all pieces of temporal information about the PoIs producing a graph with the temporal distances between each pair of PoIs. Walking distance was calculated when both PoIs were in the city centre, while the average transfer time by public transport was used in case of longer distances. We then produced a temporal knowledge base with the opening times, minimum and maximum visiting time for each PoI, and a qualitative description (''low'', ''medium'', ''high'') of the average crowding for each PoI on each day of the week. Finally, we associated with each PoI information on the services available on site (coffee shop, restaurant, . . .). For the sake of simplicity, as well as to avoid possible bias due to the specific method used to infer user preferences, in our prototype the preference values for the temporal dimensions are explicitly provided by users themselves (see Section VIII for details).

The algorithm has been run on this knowledge base using stereotypical users with different temporal preferences. The only aim of this test was the tuning of the algorithm. In particular, the test allowed us to make choices regarding the initial population and crossover and mutation strategies. Moreover, thanks to this test we observed that after 20 iterations the quality of the population (i.e., the evaluation of the best-ranked itineraries) tended to stabilize.

After the tuning, the prototype was exploited for the user evaluation that will be described in the next section.

Implementation details. Our genetic algorithm was implemented using Python. Thus, we chose Flask, together with Jinja2 template engine and standard Web technologies such as HTML5 and CSS3, for the implementation of the web interface.[1]

VIII. EVALUATION

With our evaluation, we aimed at empirically assessing our idea that users prefer itinerary recommendations which take into account their preferences on temporal dimensions (namely: minimization of transfer times, maximization of

[1]The prototype (in Italian) is available at http://fabianavernero.eu.pythonanywhere.com/.

the number of PoIs per itinerary, maximization of the heterogeneity among PoIs in a certain itinerary, inclusion of free time slots, busy hours avoidance), in addition to those regarding PoIs (including users' preference for the maximization of the number of *must see* PoIs per itinerary). To this end, we decided to carry out user-centric research, one of the three main types of general paradigms traditionally used in the evaluation of recommender systems [59]. In line with the idea that the ultimate goal of recommenders is to help users make good choices [60], user-centric research allows to gather insights about user needs and perceptions (as done for example by [3] and [19]), and is often used also to compare against baselines or alternative versions of a system (as done for example by [1] and [22]). More specifically, we compared three lists of recommendations, one generated based on the aforementioned criteria, namely:

- PP+TP: user preferences for PoIs and temporal dimensions.

and the other two generated through baseline approaches which take into account different criteria, namely:

- PP: only user preferences for PoIs,
- PP+TS: user preferences for PoIs and a standard distribution of interests for temporal dimensions.

For each recommended itinerary, we asked users to assess several aspects, namely: organisation, selection of included PoIs, and availability of free time, as well as to express an overall rating and to state how likely they would be to follow the proposed itinerary (acceptance).

Our approach assumes that user preferences for temporal dimensions and PoIs are available to the recommendation algorithm. For the purpose of our evaluation, we explicitly collected such information from users. In particular, since the automatic computation of user preferences for specific PoIs, suitable to be included in recommended itineraries, is out of the scope of this work, we decided to contextualize our evaluation in the city of Turin, which we expected to be well known to most of the people we could invite to participate in the study, and to have users directly express their liking for a selection of PoIs located in that city. Similarly, users had to explicitly express their preferences for temporal dimensions.

The evaluation was conducted online. Ethical approval for this study was obtained from the bioethical committee of the University of Turin on December 13th, 2022, with approval protocol number: 0631962, issued on December 30th, 2022.

A. METHODOLOGY

1) HYPOTHESES

We hypothesize that itinerary recommendations which take into account user preferences for contextual factors (i.e., PP+TP) receive significantly higher evaluations than itinerary recommendations which only include preferences for PoIs (i.e., PP) or non-personalized contextual factors (i.e., PP+TS), as far as the overall rating, acceptance, organisation and availability of free time aspects are concerned (H1). On the other hand, we do not expect to observe significant differences regarding the selection of PoIs, since this aspect does not depend on the inclusion of preferences for contextual factors (H2).

2) DESIGN

We adopted a within-subjects design, where the independent variable is the recommendation type, with three possible levels (PP, PP+TS, PP+TP). Thus, all participants received recommendations of all three types. Notice that, for avoiding order effects, we randomized recommendation types for each participant.

3) PARTICIPANTS

Participants were recruited using snowball sampling, a non-probability sampling technique[2] which began with a convenience sample of a few participants selected among the acquaintances of the authors. An invitation to take part in the evaluation was distributed online, through social media, mailing lists and personal messaging applications. The invitation also contained the request to forward the message to one's own contacts, so as to help the authors in finding other participants. Participants had to satisfy three criteria:

- being aged 18 or more;
- being familiar with the city of Turin;
- having read and accepted an informed consent form, through which they expressed their willingness to participate in the evaluation.

Thanks to power analysis, we determined that 119 is the minimum number of participants needed for statistically significant results with $\alpha = 0.05$, power $= 0.90$, and effect size $=$ 0.3. Our invitation was accepted by 144 people. However, only 124 of them (hence, still a large enough sample), aged from 20 to over 60 (54% in the 20-29 age range, 21% in the 30-39 age range, 14,5% in the 40-49 age range, 4,8% in the 50-59 age range and 5,6% in the 60 and over age range), 53% females, completed the whole experimental procedure.

4) APPARATUS AND MATERIAL

To collect information about the participants and to display recommended itineraries, we used the web prototype introduced in Section VII. The prototype guides participants through a step-by-step procedure (see Section VIII-A5), where each page has a specific and distinct goal, such as collecting data on a particular topic (e.g., participants' demographics) or presenting and assessing a certain recommendation approach (i.e., PP, PP+TS, or PP+TP) of itinerary recommendations. The prototype is available online and can

[2]Even though random sampling is the best way to obtain a representative sample, these strategies require a great deal of time and money. Non-probabilistic sampling is considered acceptable in HCI research [61], as well as in the user-centric evaluation of recommender systems [59].

be accessed through any browser connected to the Internet. All data were collected and stored anonymously, by assigning participants a random identifier.

5) PROCEDURE

When they accessed our online prototype, participants were presented with a welcome page which introduced the context and goals of our evaluation and gave them an overview of the steps they would go through if they agreed to participate. After having filled out the informed consent form, participants were asked to provide information on the following aspects:

- Demographics. Participants declared their age, gender and level of familiarity with the city of Turin and its PoIs.
- Preferences for PoIs. Participants explicitly expressed their preferences about the PoIs in our knowledge base, using a Likert-like scale ranging from 1 (“I’d prefer to avoid this PoI”) to 5 (“I *must see* this PoI”). Notice that PoIs were randomized for each participant, to avoid order effects.
- Preferences for the selection of PoIs. Participants expressed their level of agreement towards the following statement, using a Likert-like scale ranging from 1 (“Strongly disagree”) to 5 (“Strongly agree”): “When I plan an itinerary, I find it important to maximize the number of *must see* PoIs”.
- Preferences for temporal dimensions. Participants expressed their level of agreement towards a series of statements, each of them concerning a different temporal dimension, using a Likert-like scale ranging from 1 (“Strongly disagree”) to 5 (“Strongly agree”). All statements started with: “When I plan an itinerary, I find it important to…”, while temporal dimensions were described as follows: “minimize transfer times”, “maximize the number of PoIs to visit”, “maximize the heterogeneity among PoIs”, “include free time slots, with no pre-planned activities”, “avoid to visit very crowded PoIs”).

Notice that having users explicitly provide their preferences for specific PoIs, for PoI selection and for temporal dimensions allowed us to avoid making predictions about their possible ratings for such preferences, an aspect which is out of the scope of our research and which might have influenced users’ evaluations of the recommended itineraries. Instead, since all preferences are provided by participants themselves, they represent some sort of “ground truth” which we do not expect to play a confounding role in the evaluation of recommended itineraries. Preferences for PoIs and for temporal dimensions, which were originally expressed on a [1, 5] scale, were rescaled so that they ended up in the [0, 1] range, as explained in Section VI. Then, the genetic algorithm was run three times for each participant, once excluding temporal dimensions (PP), once taking such dimensions into account, but with predefined weights (PP+TS), equal for all participants,[3] and, finally, once considering participants’ preferences over the temporal dimensions (PP+TP). As a following step, participants were presented with three alternative itinerary recommendations (each one displayed on a separate page and randomized to avoid order effects) and were asked to assess each of them according to different aspects (see Section VIII-A6). Finally, participants were asked to provide any free comments they might like to share with us to better express their opinions about the recommendations they received; after that, they were thanked for their participation in the experiment.

TABLE 1. Participants’ evaluations (average and standard deviation) of five different aspects (overall, acceptance, organisation, POIs, free time) of itinerary recommendations.

		PP	PP+TS	PP+TP
Overall rating	Avg.	3.073	3.145	3.379
	st. dev.	0.997	0.943	0.951
Acceptance	Avg.	2.855	2.887	3.129
	st. dev.	1.167	1.053	0.987
Organisation	Avg.	3.185	3.234	3.315
	st. dev.	1.077	1.083	1.047
PoIs	Avg.	3.500	3.661	3.685
	st. dev.	1.158	0.936	0.974
Free time	Avg.	3.25	2.879	3.435
	st. dev.	1.159	1.220	1.191

6) MEASURES

For each recommended itinerary, we collected participants’ evaluations on the following aspects:

- overall evaluation (i.e., how good the recommended itinerary is overall),
- acceptance (i.e., how likely it is that the recommendation is actually accepted),
- organisation (i.e., how good the recommended itinerary is considering transfer times, number of PoIs, heterogeneity among PoIs, busy hours avoidance),
- PoIs (i.e., how good the recommended itinerary is as for the choice of PoIs to visit)
- free time (i.e., how good the recommended itinerary is as for the quantity of free time).

All evaluations were expressed using a Likert-like scale ranging from 1 (minimum) to 5 (maximum).

B. RESULTS

Descriptive statistics for all our dependent variables are reported in Table 1.[4] We can observe that recommendations generated taking into account both participants’ preferences for PoIs and temporal dimensions (PP+TP) were assessed more positively than the other two types of recommendations,

[3] Preferences for most temporal dimensions, i.e, quantity of PoIs, variety of PoIs and busy hour avoidance, were set to an intermediate score, 0.5, while preferences for transfer times and PoI’s ranking were assigned the highest possible score, 1, since they are taken into account by most itinerary recommender systems. Preferences for free time were not considered (i.e., they were set to 0).

[4] The dataset resulting from our evaluation is publicly available at the following link: https://shorturl.at/rJLY2

TABLE 2. Comparison of participants' evaluations of five different aspects (overall, acceptance, organisation, POIs, free time) between PP+TP recommendations and baseline (PP, PP+PS) recommendations - Wilcoxon test.

Aspect	Baseline	n	W	z	sign. at $p < .05$	eff. size
Overall	PP	79	1027.5	-2.7001	Y	0.304
	PP+PS	79	1167.5	-2.0159	Y	0.227
Acceptance	PP	92	1605.5	-2.0774	Y	0.217
	PP+PS	78	1120.5	-2.0919	Y	0.237
Organisation	PP	82	1473	-1.0563	N	n.a.
	PP+PS	84	1676.5	-0.4839	N	n.a.
PoIs	PP	76	1160	-1.5687	N	n.a.
	PP+PS	66	1051	-0.3482	N	n.a.
Free time	PP	78	1244.5	-1.4743	N	n.a.
	PP+PS	72	547.5	-4.3014	Y	0.507

as far as all aspects are concerned. PP+TS recommendations also scored slightly better than PP, apart for the "free time" aspect, where they obtained the lowest score. Notice that both PP and PP+TS recommendations did not include any free time slots; however, PP recommendations were judged more positively than PP+TS about this aspect, with 56% participants declaring to be satisfied in the former condition (a result quite close to that obtained by PP+TP recommendations, i.e., 63%), and only 20% in the latter case.

Overall, these results are in line with those we obtained in the preliminary study described in [16], where we assessed the same three types of itineraries with a small sample of 20 users. In fact, we also found that PP+TP recommendations were evaluated more positively, and obtained an average score of 3.4 (st. dev: 0.94) for the "overall evaluation" aspect, which was higher than the scores obtained by PP+TS (avg: 3.1, st. dev: 1.07) and PP (avg: 3.2, st. dev: 0.89) recommendations. Differently from the current study, however, we did not assess more specific aspects on that occasion.

In order to validate our hypotheses (see Section VIII-A1), we compared participants' ratings for PP+TP recommendations with ratings for PP+PS and PP recommendations, so as to ascertain whether the observed differences are statistically significant. Since the Kolmogorov-Smirnov test showed that all dependent variables were *not* normally distributed, we chose to apply the Wilcoxon test[5] to carry out pairwise comparisons. Results are reported in Table 2.

1) H1: OVERALL ASSESSMENT AND CONTEXT-RELATED ASPECTS

As for the "overall evaluation" aspect, the ratings participants assigned to PP+TP recommendations (avg: 3.379) are significantly higher than the scores obtained by PP (avg: 3.073, $W = 1027.5$, $p < .05$, $n = 79$) and PP+PS (avg: 3.145, $W = 1167.5$, $p < .05$, $n = 79$) recommendations, with a medium and a small effect size, respectively, according to Cohen's classification.

Similarly, the differences in participants' evaluations are significant also for the "acceptance" aspect, with the ratings assigned to PP+TP recommendations (avg: 3.129) being significantly better than those assigned to both PP (avg: 2.855, $W = 1605.5$, $p < .05$, $n = 92$) and PP+TS (avg: 2.887, $W = 1120.5$, $p < .05$, $n = 78$), with a small effect size in both cases.

On the contrary, as for the "organisation" aspect, the observed differences in participants' ratings are too small to be significant, which is a bit surprising, since this facet is connected to several contextual dimensions and, as such, should better match participants' preferences when these are explicitly taken into account, as in PP+TP recommendations. To better understand this unexpected finding, we adopted a qualitative approach, inspired by thematic analysis [62], to examine the free comments provided by participants at the end of the test. After a first read-through of participants' answers, where we tried to identify potential categories for them based on emerging and recurrent topics in their description of issues with the recommended itineraries, we systematically re-examined all the comments, pulling out key points and labelling them with the previously defined categories. Out of 57 comments, only 4 were completely positive or non relevant, while the others were found to deal with problems related to one or more of the categories listed in Table 3. Interestingly, the most popular category regards transfers (21 comments) and, in particular, transfer distances (15 comments): in fact, several participants pointed out that transfers between consecutive PoIs were not efficient, since recommended itineraries suggested to visit PoIs which are far from each other one after the other (P148[6]: "Points of interest are not always close to each other. When I'm on holiday, I prefer to see as many things as possible without having to move too much."), to move back and forth between different areas (P107: "There are often long transfers: for example, from Turin to Venaria[7] and then back to the city center, near the starting point"; P148: "Going back and forth is annoying"; P154: "I'd prefer to visit tourist attractions in the outskirts in a dedicated morning or afternoon, thus moving from/to Turin only once") or to visit nearby PoIs on different days (P125: "Places very close to each other were scheduled on different days or at different times, which is not very efficient"). This is probably due to the fact that the genetic algorithm was programmed to optimize temporal dimensions globally (i.e., considering the itinerary as a whole) and not locally (i.e., considering every single transfer between two consecutive PoIs), as explained in Section VI-B. Based on participants' comments, however, this approach does not seem to completely satisfy user needs, and/or to be in line with their expectations, which can probably explain the lower than expected ratings on the "organisation" aspect. In addition, a few participants (6 comments) also complained

[5]The Wilcoxon test is a nonparametric measure which evaluates the difference between two conditions in a within-subjects design.

[6]Participants will be identified through their anonymous codes.

[7]Venaria Reale is a municipality located about 8 kilometres northwest of Turin.

TABLE 3. Categories (i.e., recurring topics) identified during the the qualitative analysis of free comments, with their respective comment count.

Categories	Count
Transfers	21
Transfer distances	*15*
Estimated transfer times	*5*
PoI selection	18
Visit duration	12
Free time	6
Miscellanea	4
PoI heterogeneity	3
Number of PoIs	2

about estimated transfer times, which they deemed to be unrealistic (P111: "Transfer times to get to places situated in the outskirts are too short"; P145: "Estimated transfer times should be longer: tourists might want to window shop or quietly amble down the streets. Moreover, streets might be crowded, for example on the weekends."), for example because they failed to take into account the time needed to hunt for parking spaces (P168: "I don't think transfer times consider the time needed to park in the city center."). Differently from the previous comments on transfer distances, these complaints draw attention to an issue with the input data, rather than with the recommendation algorithm, and underline the importance of accurate knowledge bases.

Finally, as for the "free time" aspect, PP+TP recommendations (avg: 3.435) obtained significantly higher ratings only in comparison with PP+TS (avg: 2.879, $W = 547.5$, $p < .05$, $n = 72$). Notice that ratings for PP (avg: 3.25) were unexpectedly high, and they were even significantly higher than those for PP+TS recommendations (avg: 2.879, $W = 747$, $p < .05$, $n = 73$), which is quite surprising, considering that both these types included no free time slots, as mentioned before. One possible explanation is that a few participants who have received PP as the first type of recommendations to evaluate might have considered the fact that itineraries end quite early around dinner time (a feature which is the same for all types of recommendations) when assessing the amount of free time, only to notice at later steps that free time slots, when available, are scheduled during the itinerary and are explicitly labelled as such. Considering that this presentation choice was mentioned in the instructions to participants, but the absence of free time slots was not highlighted on a per-itinerary basis, we are prone to interpret this unexpected result as the consequence of our design choices, which might have confused a few participants, rather than as the expression of real preferences towards PP versus PP+TS recommendations.

To sum up, these results allow us to partially confirm our first hypothesis (H1): in fact, itinerary recommendations which take into account user preferences for contextual factors are more appreciated than the other two types of recommendations as far as the overall rating and acceptance aspects are concerned. However, we found no (or only partial) significant differences as for the organisation and availability of free time aspects, which calls for further tuning of both our genetic algorithm and the prototype interface.

2) H2: SELECTION OF POIS

As for the "PoI" aspect, which aimed at measuring how good the recommended itinerary is as for the choice of suggested PoIs, we found no significant differences, neither between PP+TP and PP, nor between PP+TP and PP+TS recommendations. This result is in line with our expectations since the procedure for the selection of PoIs does not vary across the three recommendation types. Hence, we can confirm our second hypothesis (H2).

IX. LIMITATIONS

Our work has some limitations.

As for our experimental evaluation, we acknowledge that, while our sample is well-balanced gender-wise and covers a wide range of ages (20->60), it does not accurately reflect the features of the whole population. For example, most participants fall in the 20-29 age range, and they are all already familiar with the city of Turin (a requirement set to make the initial evaluation of PoIs easier for them). A more diverse sample would help to improve the external validity of our study. In addition, we are aware that participants in a controlled experiment can in general have different goals and motivations if compared with real users, which may have impacted the perception of the recommendations provided in our evaluation.

Moreover, our approach takes into account only positive user preferences, both as regards the PoIs to visit and as regards temporal preferences. It is very important to consider also negative preferences, which can play an important role in the recommendation process [63], [64] and are particularly relevant in the tourist domain (e.g., for specifying PoIs to be avoided or negative temporal preferences such as avoiding long transfers between PoIs). In future work, we plan to extend it to take into account also negative preferences in the generation of the itinerary.

Also, we considered only a few heuristics for evaluating itineraries concerning their five temporal dimensions. We chose those that are simpler to compute and that seemed to provide better results during the tuning phase. However, it could be interesting to use a wider set of heuristics. In future work, we could also experiment with a different variety of heuristics.

Finally, as of now our approach for generating an itinerary is static and does not take into account the possibility of dynamically changing the itinerary on the fly. Tourists may change their preferences during the different phases of their trip. Thus, another possible future direction of the work is to consider the possibility of dynamically changing the itinerary on the fly.

X. CONCLUSION

In this paper we proposed an approach for generating personalized itineraries that take into account a variety of user preferences, focusing in particular on temporal ones. We chose to design and implement the system as a service that can be coupled with any recommender suggesting a tourist the PoIs that are of interest for her/him. Indeed our system requires a ranked list of PoIs as its input. This decision had a tremendous impact in the evaluation process since it allowed us to evaluate the itinerary planner independently of the problem of recommending PoIs; in particular in the evaluation we assumed that the ranked list of PoIs is provided by users themselves and thus the results we obtained in the evaluation of the suitability of the proposed itineraries is independent of that list (and of the criteria and approach to generate it).

In our approach, we took into account time in a thorough way, considering both objective (such as time distances between PoIs, time for each visit, ...) and subjective temporal information (user temporal preferences). This is a significant difference and improvement with respect to the literature.

Genetic algorithms proved to be an interesting solution for our optimization problem, where user temporal preferences play a peculiar role in the evaluation of the fitness of an itinerary. The algorithm we adopted deals in a natural way with the combinatorial nature of the problem allowing a sort of parallel exploration of alternative itineraries. In this sense we claim that, for this type of problem, genetic algorithms are easier to adopt than other traditional approaches to temporal planning [65]. Moreover they have the advantage of being any-time. The genetic algorithm has then been coupled with constraint satisfaction to check objective temporal constraints.

Two evaluation processes have been carried out: the first, with stereotypical users, to tune the algorithm, and the second, with actual users, to evaluate the suitability of the itineraries being proposed. The latter evaluation provided interesting results and also insights on potential extensions of the approach. First of all tuning has been performed on a single application and should be further investigated using other applications. Tuning led us to make some choices in the genetic algorithm (concerning, e.g., crossover and mutation), selecting those that led to the best performance. Many other approaches to perform crossover and mutations could be considered and compared to the one we chose.

Furthermore, since travelling is a group activity, it is important to consider how to generate group recommendations [66], with particular attention to children. In fact, when organising activities, one must always take into account the possible presence of children, whose needs and preferences influence the travel experience.

ACKNOWLEDGMENT

(Federica Cena, Luca Console, Marta Micheli, and Fabiana Vernero contributed equally to this work.)

REFERENCES

[1] K. Taylor, K. H. Lim, and J. Chan, "Travel itinerary recommendations with must-see points-of-interest," in *Proc. Companion Web Conf.*, 2018, pp. 1198–1205.

[2] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Model. User-Adapted Interact.*, vol. 12, no. 4, pp. 331–370, 2002.

[3] M. Tenemaza, S. Lujan-Mora, A. De Antonio, and J. Ramirez, "Improving itinerary recommendations for tourists through metaheuristic algorithms: An optimization proposal," *IEEE Access*, vol. 8, pp. 79003–79023, 2020.

[4] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, "Personalized itinerary recommendation with queuing time awareness," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2017, pp. 325–334.

[5] P. L. Pearce, "Tourists' perception of time: Directions for design," *Ann. Tourism Res.*, vol. 83, Jul. 2020, Art. no. 102932.

[6] L. Botti, N. Peypoch, and B. Solonandrasana, "Time and tourism attraction," *Tourism Manage.*, vol. 29, no. 3, pp. 594–596, Jun. 2008.

[7] A. Zalatan, "The determinants of planning time in vacation travel," *Tourism Manage.*, vol. 17, no. 2, pp. 123–131, Mar. 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0261517795001158

[8] C. Zhang, H. Liang, K. Wang, and J. Sun, "Personalized trip recommendation with poi availability and uncertain traveling time," in *Proc. 24th ACM Int. Conf. Inf. Knowl. Manage.*, 2015, pp. 911–920.

[9] C. Zhang, H. Liang, and K. Wang, "Trip recommendation meets real-world constraints: POI availability, diversity, and traveling time uncertainty," *ACM Trans. Inf. Syst.*, vol. 35, no. 1, pp. 1–28, Jan. 2017.

[10] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized tour recommendation based on user interests and points of interest visit durations," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 1–7.

[11] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," *Knowl. Inf. Syst.*, vol. 54, no. 2, pp. 375–406, Feb. 2018.

[12] J.-H. Chen, K.-M. Chao, and N. Shah, "Hybrid recommendation system for tourism," in *Proc. IEEE 10th Int. Conf. e-Bus. Eng.*, Nov. 2013, pp. 156–161.

[13] W. Zheng, Z. Liao, and J. Qin, "Using a four-step heuristic algorithm to design personalized day tour route within a tourist attraction," *Tourism Manage.*, vol. 62, pp. 335–349, Oct. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0261517717301061

[14] B. S. Wibowo and M. Handayani, "A genetic algorithm for generating travel itinerary recommendation with restaurant selection," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2018, pp. 427–431.

[15] P. Yochum, L. Chang, T. Gu, and M. Zhu, "An adaptive genetic algorithm for personalized itinerary planning," *IEEE Access*, vol. 8, pp. 88147–88157, 2020.

[16] F. Cena, L. Console, M. Micheli, and F. Vernero, "Combining genetic algorithms and temporal constraint satisfaction for recommending personalized tourist itineraries," in *Proc. Int. Conf. Italian Assoc. Artif. Intell. (AIxIA)*, vol. 14318, R. Basili, D. Lembo, C. Limongelli, and A. Orlandini, Eds., Rome, Italy. Cham, Switzerland: Springer, Nov. 2023, pp. 441–452, doi: 10.1007/978-3-031-47546-7_30.

[17] B. Ludwig, B. Zenker, and J. Schrader, "Recommendation of personalized routes with public transport connections," in *Intelligent Interactive Assistance and Mobile Multimedia Computing*, D. Tavangarian, T. Kirste, D. Timmermann, U. Lucke, and D. Versick, Eds., Berlin, Germany: Springer, 2009, pp. 97–107.

[18] K.-P. Chang, L.-Y. Wei, M.-Y. Yeh, and W.-C. Peng, "Discovering personalized routes from trajectories," in *Proc. 3rd ACM SIGSPATIAL Int. Workshop Location-Based Social Netw.* New York, NY, USA: ACM, Nov. 2011, pp. 33–40.

[19] D. Quercia, R. Schifanella, and L. M. Aiello, "The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city," in *Proc. 25th ACM Conf. Hypertext Social Media*. New York, NY, USA: ACM, Sep. 2014, pp. 116–125.

[20] S. Comai, E. D. Bernardi, M. Matteucci, and F. Salice, "Maps for easy paths (MEP): Enriching maps with accessible paths using mep traces," in *Smart Objects and Technologies for Social Good*, O. Gaggi, P. Manzoni, C. Palazzi, A. Bujari, and J. M. Marquez-Barja, Eds., Cham, Switzerland: Springer, 2017, pp. 254–263.

[21] J. Kim, M. Cha, and T. Sandholm, "Socroutes: Safe routes based on tweet sentiments," in *Proc. 23rd Int. Conf. World Wide Web*. New York, NY, USA: ACM, 2014, pp. 179–182.

[22] M. De Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu, "Automatic construction of travel itineraries using social breadcrumbs," in *Proc. 21st ACM Conf. Hypertext hypermedia*, Jun. 2010, pp. 35–44.

[23] H.-P. Hsieh, C.-T. Li, and S.-D. Lin, "TripRec: Recommending trip routes from large scale check-in data," in *Proc. 21st Int. Conf. World Wide Web*, Apr. 2012, pp. 529–530.
[24] X. Li, "Multi-day and multi-stay travel planning using geo-tagged photos," in *Proc. 2nd ACM SIGSPATIAL Int. Workshop Crowdsourced Volunteered Geographic Inf.*, Nov. 2013, pp. 1–8.
[25] I. Refanidis, C. Emmanouilidis, I. Sakellariou, A. Alexiadis, R. A. Koutsiamanis, K. Agnantis, A. Tasidou, F. Kokkoras, and P. S. Efraimidis, "MYVISITPLANNER GR: Personalized itinerary planning system for tourism," in *Proc. Hellenic Conf. Artif. Intell.* Cham, Switzerland: Springer, 2014, pp. 615–629.
[26] L. Chen, L. Zhang, S. Cao, Z. Wu, and J. Cao, "Personalized itinerary recommendation: Deep and collaborative learning with textual information," *Expert Syst. Appl.*, vol. 144, Apr. 2020, Art. no. 113070.
[27] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Smart itinerary recommendation based on user-generated GPS trajectories," in *Proc. Int. Conf. Ubiquitous Intell. Comput.* Cham, Switzerland: Springer, 2010, pp. 19–34.
[28] A. Fogli, A. Micarelli, and G. Sansonetti, "Enhancing itinerary recommendation with linked open data," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2018, pp. 32–39.
[29] G. Cai, K. Lee, and I. Lee, "Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos," *Expert Syst. Appl.*, vol. 94, pp. 32–40, Mar. 2018.
[30] M. De Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu, "Constructing travel itineraries from tagged geo-temporal breadcrumbs," in *Proc. 19th Int. Conf. World Wide Web*, Apr. 2010, pp. 1083–1084.
[31] P. D. Bitonto, F. D. Tria, M. Laterza, T. Roselli, V. Rossano, and F. Tangorra, "Automated generation of itineraries in recommender systems for tourism," in *Proc. 10th Int. Conf. Current Trends Web Eng.* Berlin, Germany: Springer, 2010, pp. 498–508.
[32] A. Gionis, T. Lappas, K. Pelechrinis, and E. Terzi, "Customized tour recommendations in urban areas," in *Proc. 7th ACM Int. Conf. Web Search Data Mining*, Feb. 2014, pp. 313–322.
[33] I. Garcia, L. Sebastia, and E. Onaindia, "On the design of individual and group recommender systems for tourism," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7683–7692, Jun. 2011.
[34] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Towards next generation touring: Personalized group tours," *Proc. Int. Conf. Automated Planning Scheduling*, vol. 26, pp. 412–420, Mar. 2016.
[35] H.-P. Hsieh, C.-T. Li, and S.-D. Lin, "Measuring and recommending time-sensitive routes from location-based data," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, pp. 1–27, Jul. 2014.
[36] C. Biancalana, F. Gasparetti, A. Micarelli, A. Miola, and G. Sansonetti, "Context-aware movie recommendation based on signal processing and machine learning," in *Proc. 2nd Challenge Context-Aware Movie Recommendation*. New York, NY, USA: ACM, Oct. 2011, pp. 5–10.
[37] G. Arru, D. Feltoni Gurini, F. Gasparetti, A. Micarelli, and G. Sansonetti, "Signal-based user recommendation on Twitter," in *Proc. 22nd Int. Conf. World Wide Web*. New York, NY, USA: ACM, May 2013, pp. 941–944.
[38] G. Sansonetti, D. F. Gurini, F. Gasparetti, and A. Micarelli, "Dynamic social recommendation," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*. New York, NY, USA: ACM, Jul. 2017, pp. 943–947.
[39] I. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "Where shall we go today? Planning touristic tours with tripbuilder," in *Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage.*, 2013, pp. 757–762.
[40] I. R. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "On planning sightseeing tours with TripBuilder," *Inf. Process. Manage.*, vol. 51, no. 2, pp. 1–15, Mar. 2015.
[41] B. Gavish and S. C. Graves, "The travelling salesman problem and related problems," Oper. Res. Center, MIT, Work. Paper GR-078-78, 1978.
[42] F. Cena, L. Console, S. Likavec, M. Micheli, and F. Vernero, "How personality traits can be used to shape itinerary factors in recommender systems for young travellers," *IEEE Access*, vol. 11, pp. 61968–61985, 2023.
[43] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Social itinerary recommendation from user-generated digital trails," *Pers. Ubiquitous Comput.*, vol. 16, no. 5, pp. 469–484, Jun. 2012.
[44] J. Zhou, Y. Gu, and W. Lin, "Complementing travel itinerary recommendation using location-based social networks," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Advanced Trusted Comput., Scalable Comput. Commun., Cloud Big Data Computing, Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Sep. 2019, pp. 1876–1881.
[45] H. Ji, W. Zheng, X. Zhuang, and Z. Lin, "Explore for a day? Generating personalized itineraries that fit spatial heterogeneity of tourist attractions," *Inf. Manage.*, vol. 58, no. 8, Dec. 2021, Art. no. 103557.
[46] E. Erbil and W. Wörndl, "Personalization of multi-day round trip itineraries according to travelers' preferences," in *Proc. ENTER eTourism Conf.* Cham, Switzerland: Springer, Jan. 2022, pp. 187–199.
[47] S. Halder, K. H. Lim, J. Chan, and X. Zhang, "Efficient itinerary recommendation via personalized POI selection and pruning," *Knowl. Inf. Syst.*, vol. 64, no. 4, pp. 963–993, Apr. 2022.
[48] J. Karbowska-Chilinska and P. Zabielski, *Genetic Algorithm Solving Orienteering Problem With Time Windows*, vol. 240. Cham, Switzerland: Springer, Sep. 2013.
[49] C. Changdar, G. S. Mahapatra, and R. K. Pal, "An efficient genetic algorithm for multi-objective solid travelling salesman problem under fuzziness," *Swarm Evol. Comput.*, vol. 15, pp. 27–37, Apr. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210650213000679
[50] C. Changdar, R. K. Pal, and G. S. Mahapatra, "A genetic ant colony optimization based algorithm for solid multiple travelling salesmen problem in fuzzy rough environment," *Soft Comput.*, vol. 21, no. 16, pp. 4661–4675, Aug. 2017, doi: 10.1007/s00500-016-2075-4.
[51] N. N. Qomariyah and D. Kazakov, "A genetic-based pairwise trip planner recommender system," *J. Big Data*, vol. 8, p. 77, May 2021.
[52] F. Ghobadi, A. Divsalar, H. Jandaghi, and R. Barzegar Nozari, "An integrated recommender system for multi-day tourist itinerary," *Appl. Soft Comput.*, vol. 149, Dec. 2023, Art. no. 110942. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494623009602
[53] H. Jandaghi, A. Divsalar, and S. Emami, "The categorized orienteering problem with count-dependent profits," *Appl. Soft Comput.*, vol. 113, Dec. 2021, Art. no. 107962. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S156849462100884X
[54] R. Dechter, I. Meiri, and J. Pearl, "Temporal constraint networks," *Artif. Intell.*, vol. 49, nos. 1–3, pp. 61–95, May 1991, doi: 10.1016/0004-3702(91)90006-6.
[55] V. Brusoni, L. Console, and P. Terenziani, "On the computational complexity of querying bounds on differences constraints," *Artif. Intell.*, vol. 74, no. 2, pp. 367–379, Apr. 1995.
[56] R. R. Sharapov, "Genetic algorithms: Basic ideas, variants and analysis," in *Vision Systems*, G. Obinata and A. Dutta, Eds., Rijeka, Croatia: IntechOpen, 2007, ch. 21, doi: 10.5772/4971.
[57] M. S. Malik and M. S. Wadhwa, "Preventing premature convergence in genetic algorithm using DGCA and elitist technique," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 6, pp. 1–9, 2014.
[58] A. Hassanat, K. Almohammadi, E. Alkafaween, E. Abunawas, A. Hammouri, and V. B. S. Prasath, "Choosing mutation and crossover ratios for genetic algorithms—A review with a new dynamic approach," *Information*, vol. 10, no. 12, p. 390, Dec. 2019.
[59] D. Jannach, "Evaluating conversational recommender systems: A landscape of research," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 2365–2400, Jul. 2022, doi: 10.1007/s10462-022-10229-x.
[60] A. Jameson, M. C. Willemsen, A. Felfernig, M. de Gemmis, P. Lops, G. Semeraro, and L. Chen, "Human decision making and recommender systems," in *Recommender Systems Handbook*. Cham, Switzerland: Springer, 2015, pp. 611–648.
[61] J. Lazar, J. H. Feng, and H. Hochheiser, "Surveys," in *Research Methods in Human Computer Interaction*, 2nd ed., J. Lazar, J. H. Feng, and H. Hochheiser, Eds., Boston, MA, USA: Morgan Kaufmann, 2017, pp. 105–133. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128053904000054
[62] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Res. Psychol.*, vol. 3, no. 2, pp. 77–101, Jan. 2006.
[63] F. Cena, L. Console, and F. Vernero, "How to deal with negative preferences in recommender systems: A theoretical framework," *J. Intell. Inf. Syst.*, vol. 60, no. 1, pp. 23–47, Feb. 2023.
[64] F. Cena, L. Console, and F. Vernero, "Logical foundations of knowledge-based recommender systems: A unifying spectrum of alternatives," *Inf. Sci.*, vol. 546, pp. 60–73, Feb. 2021, doi: 10.1016/j.ins.2020.07.075.

[65] M. Ghallab, D. S. Nau, and P. Traverso, *Automated Planning Acting*. Cambridge, U.K.: Cambridge Univ. Press, 2016. [Online]. Available: http://www.cambridge.org/de/academic/subjects/computer-science/artificial-intelligence-and-natural-language-processing/automated-planning-and-acting?format=HB

[66] A. Jameson and B. Smyth, "Recommendation to groups," in *The Adaptive Web: Methods and Strategies of Web Personalization*. Berlin, Germany: Springer, 2007, pp. 596–627.

FEDERICA CENA is an Associate Professor with the Computer Science Department, University of Turin. She works on the intersection of artificial intelligence and human–computer interaction. In the last years, she is mainly devoted to studying user modeling and personalization for cognitive disabilities and frailty. She is the author of more than 150 scientific publications at conferences and international journals.

LUCA CONSOLE received the Ph.D. degree in computer science in 1991. He is currently a Full Professor of computer science with the University of Turin. He has co-authored the collection of seminal works on model-based diagnosis and published over 150 papers in international journals and conferences. He has coordinated several projects and research contracts from several national industries. His research interests include different areas of artificial intelligence, ranging from logical foundation of reasoning to model-based reasoning and diagnosis, adaptive systems, smart physical objects, and intelligent interaction. He served in the program committee for many major conferences and has been an invited speaker in some of them, including IJCAI 1999.

MARTA MICHELI is currently pursuing the Ph.D. degree in computer science with the University of Turin, Italy. She has an interdisciplinary background. Her research interests include user modeling, with a particular focus on detecting and integrating psychological information into personalization systems.

FABIANA VERNERO received the Ph.D. degree in computer science. She is currently an Associate Professor with the Department of Computer Science, University of Turin, where she works in the human–computer interaction area. Her research interests include intelligent user interfaces, recommender systems, and persuasive technologies. In the past, she founded a startup in the mobile app industry.

Open Access funding provided by 'Università degli Studi di Torino' within the CRUI CARE Agreement