J Vis
https://doi.org/10.1007/s12650-022-00884-1

REGULAR PAPER

Zikun Deng · Di Weng · Yingcai Wu

You are experienced: interactive tour planning with crowdsourcing tour data from web

Received: 22 July 2022 / Accepted: 9 September 2022

Abstract Planning an ideal tour schedule is a tedious process, where the attractions to visit and the order of visits need to be carefully determined. In this paper, we propose a novel interactive approach for tour planning. We first extract prior tourists' experiences from the crowdsourcing tour data on the Web using frequent substring mining. We then design and implement a planning tool equipped with interactive visualizations, enabling users to learn the extracted experiences and plan their own tours. Our approach is evaluated with two usage scenarios on real-world tour data in two cities. Compared with previous methods, our approach strikes a balance between efficiency and reliability. On the one hand, we support the interactive manipulation of attraction sequence (i.e., multiple attractions at a time), thereby ensuring efficiency. On the other hand, we keep humans in the loop of the tour planning process via interactive visualizations. This paper shows the value of tour data published by tourists on the Web for personalized tour planning.

Keywords Interactive tour planning · Geospatial data visualization · Mobility visualization

1 Introduction

Tourism has become an important and popular way of pastime and entertainment in people's modern life. Tour planning is a challenging task. To determine an ideal tour schedule (i.e., the attractions to visit and the order of visits), tourists or practitioners (e.g., tourist guides) need to consider attractions' popularities, travel experiences (Nomiyama et al. 2018), visit times, etc.

Many interactive tools have been developed to assist in planning tours. The strategies of these tools can be categorized into *end-to-end* (Yahi et al. 2015; Kurata and Hara 2014; Lim et al. 2016; Inspirock 2022) and *step-by-step* (Nomiyama et al. 2018; Travelchime Inc 2022; Tripadvisor Inc 2022). The *end-to-end* strategy (Fig. 1A) allows users to specify their individual preferences, such as the tour duration, tightness, and preferred POI categories (e.g., shopping). Then, automatic optimization algorithms recommend candidate tours that satisfy users' preferences as much as possible. In this way, users can quickly obtain

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12650-022-00884-1.

Z. Deng · Y. Wu
State Key Lab of CAD & CG, Zhejiang University, Hangzhou, China
E-mail: zikun_rain@zju.edu.cn

Y. Wu
E-mail: ycwu@zju.edu.cn

D. Weng (✉)
Microsoft Research Asia, Beijing, China
E-mail: diweng@microsoft.com

Published online: 26 October 2022

candidate tours via simple interactions. However, users still need to evaluate the recommended routes because they cannot always accurately externalize their preferences and fully trust the black-box recommendation, which leads to a reliability issue. The *step-by-step* strategy (Fig. 1C) presents attractions together with their context and supports users to construct tours by adding the attractions of their interests iteratively and manually. During the process, users have to carefully evaluate each added attraction and its arrangement by investigating contextual information. Compared with the end-to-end strategy, it is more flexible and reliable but indeed a tedious and time-consuming process, leading to an efficiency issue.

The aforementioned limitations observed in the existing tools motivate us to propose a new interactive approach for tour planning to feature a balance between efficiency and reliability.

The dilemma is that reliable tour planning inevitably demands a time-consuming investigation of attractions with respect to their contents (e.g., what to enjoy) and spatiotemporal relationships with each other (e.g., what is the next attraction). Fortunately, with the rise of tourism and the development of the Internet, more and more tourists voluntarily share their travelogues with itineraries on tourism web forums. Specifically, an itinerary records several attraction sequences (denoted as *tours*) that a tourist or a group of tourists visits on different days. We call these attraction sequences crowdsourcing tour data. They provide an unprecedented opportunity for tourists or practitioners to learn the previous tourists' experiences and then plan new tours efficiently and reliably, rather than starting from knowing nothing.

This study attempts to propose a novel approach for tour planning by leveraging the tour data publicly shared on the web with interactive visualizations. Notably, our approach enables users to visually analyze the popular attraction subsequences (denoted as *subtour*) from these tour data and interactively manipulate them for tour planning (Fig. 1B).

However, two major challenges are posed. **(1) Scalable and clear presentation** of crowdsourcing tour data. Many popular subtours, such as attractions "$A \rightarrow B \rightarrow C$", attractions "$B \rightarrow C$", and attractions "$C \rightarrow B \rightarrow D$," exist among the crowdsourcing tour data. They may overlap with each other or even contain the opposite order of attractions. Users need to determine their favorite by exploring these popular subtours and conducting analyses, which demands a scalable and intuitive visualization. **(2) Interactive planning** of an ideal tour. The interactions of the previous methods (whether it is end-to-end or step-by-step) focus on individual attractions. In contrast, our tool allows users to plan their ideal tours based on the (sub) tours of others, which have not been studied before. Thus, new interactive methods are required to support flexible manipulation of the tours users are interested in.

In this study, we first gather user requirements that describe how users want to plan ideal tours using crowdsourcing tour data on the Web. Based on the gathered requirements, we propose a novel tool planning framework. We also accordingly develop cT^3, a **C**rowdsourcing **T**our-based **T**our planning **T**ool. cT^3 strikes a balance between end-to-end and step-by-step strategies by introducing subtours. cT^3 addresses the aforementioned challenges as follows. (1) To address the first challenge, we tightly combine a computational method and visualizations. In particular, we identify popular subtours using a frequent substring mining technique and present them intuitively and compactly in a scalable manner based on their inherent parent–children relationships. (2) To address the second challenge, we propose a set of useful interactions to assist users in aligning, comparing, and connecting (sub) tours, thereby enabling them to derive their ideal tours interactively. The subtour concept supports planning multiple orderly attractions at once and the visualizations keep human in the loop, thereby ensuring efficiency while pursuing reliability. Finally, we present two usage scenarios in two different cities, which show the effectiveness and usability of our approach. In sum, our contributions are as follows:

- A study showing the value of tour data on the Web for tour planning, which will promote tourism forums or related websites;
- A tool planning framework and its implementation, a novel interactive tour planning tool, cT^3. To the best of our knowledge, it is the first attempt that enables users to learn the prior experiences of other tourists and plan their ideal tours efficiently and reliably;
- Two example usage scenarios on real-world data in two cities.

2 Related work

This section presents related studies in three categories: tour planning tool, tour data visualization, and urban visual analytics.

2.1 Tool planning tool

We focus on planning tours within a few days, not the next attraction (Contractor et al. 2021; Wang et al. 2020) or the next city (Dadoun et al. 2019). Many interactive tools have been developed to assist tourists in planning their tours and can be divided into two categories based on the adopted strategies: end-to-end and step-by-step.

The end-to-end tools build upon user interfaces and optimization algorithms. User interfaces allow tourists to specify their individual preferences as constraints. Then, tour planning problems are formulated as optimization problems to recommend tours that satisfy tourists' preferences as much as possible. Many algorithms (Lim et al. 2019) have been developed to do that by considering various aspects, such as peak season or not (Dunstall et al. 2003; Yim et al. 2004), tourist group (Kinoshita and Yokokishizawa 2008), transportation mode (Brennan and Meier 2007), hotel services (Silamai et al. 2017), Must-See POI (Taylor et al. 2018), route attractiveness (Herzog et al. 2019), user's current location and spare time Kurashima et al. (2010). For example, with PersTour (Lim et al. 2016; Kurata and Hara 2014), users can specify the tour duration, tightness, and preferred POI categories (e.g., shopping) and then obtain a recommended tour. Some tools, such as Aurigo (Yahi et al. 2015) and CT-Planner4 (Kurata and Hara 2014), allowed users to further refine the recommended tours on demand by adding or deleting some attractions. Using end-to-end tour planning tools, tourists can quickly obtain candidate tours with simple interactions. Nevertheless, careful evaluation of the recommended routes and even exploratory refinement are still required because tourists cannot always accurately externalize their preferences and fully trust the black-box recommendation.

The step-by-step tools do not require complex computational models. They display comprehensive and useful information about attractions and provide friendly user interactions. Tourists can construct ideal tours by adding the attractions of their interests iteratively and manually. For example, Aurigo (Yahi et al. 2015) adopted a circular lens called Pop Radius to reveal the neighboring attractions of the focus attraction and thereby facilitated step-by-step tour construction. Some commercial softwares, such as GoogleTravel (2022), Tripadvisor (2022), and Travelchime Inc (2022), also employed such a step-by-step strategy for tour planning. For instance, in GoogleTravel (2022), every attraction has several beautiful photographs, comments by tourists, and scores. Their geographic locations are encoded in a geographic map. Using step-by-step tour planning tools, tourists have carefully evaluated every added attraction by investigating their contextual information and then arranged them. However, simply based on individual attractions without explicit relationships between them, it is difficult to assess and determine their order. Compared with the

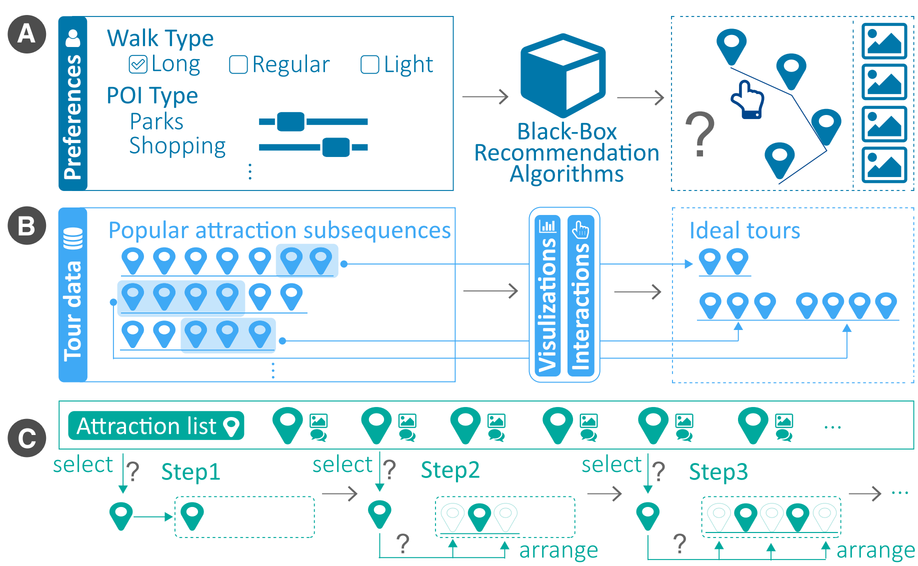

Fig. 1 Three tour planning strategies. **A** End-to-end: a user specifies his preferences and a black-box model recommends a tour; **B** our proposed tour-based: a user selects attraction sequences based on the visualization of previous tourists' tours and then constructs his tour interactively; **C** step-by-step: a user iteratively selects one attraction of his interest and arranges it into his tour. Question marks indicate that users may be in trouble

end-to-end strategy, the step-by-step strategy is a more flexible and reliable manner but indeed a more tedious and time-consuming process.

By contrast, we attempt to compromise between the two strategies. We first apply a computational model on crowdsourcing tour data shared on the Web to identify popular attraction sequences called subtours, rather than a few complete recommended tours in end-to-end strategy. Afterward, we design a set of visualizations and interactions. Users can reliably and flexibly manipulate these subtours (e.g., merge them as ideal tours). Manipulating several attractions of a subtour at once avoids tedious step-by-step processes.

2.2 Tour data visualization

Tour data describe tourists' behaviors in the cities. It can be roughly divided into two categories, namely materials (e.g., blogs (Sharda and Ponnada 2008) and photographs (Claudio and Yoon 2014; Thudt et al. 2016)) and movement records (e.g., how tourists move). In our study, we focus more on movement records.

Numerous visualization techniques have been proposed for movement data and can be divided into three categories (Andrienko et al. 2013), namely direct depiction (Klein et al. 2021), summarization (Liu et al. 2021a), and pattern extraction (Deng et al. 2020; Bai et al. 2021). For example, Kádár and Gede (2013) directly depicted the tourists' movement records as dots to measure their activity in urban spaces. Lee and Tsou (2018) directly depict tourists' trajectories with lines on the map. Hu et al. (2019) summarized massive and noisy tourist movement data with a graph representation. Chen et al. (2016) extracted movement patterns from sparsely sampled tourist movement data. In this paper, we adopt a pattern extraction strategy. In particular, we first identify the popular attraction sequences from the crowdsourcing tour data and then visually depict them for tour planning.

2.3 Urban visual analytics

Visual analytics (Andrienko et al. 2021; Deng et al. 2023) enables urban practitioners to explore and analyze urban data via interactive spatiotemporal visualizations (Liu et al. 2023; Wang et al. 2021a; Zheng et al. 2021; Wei et al. 2021). Urban visual analytics can be divided into urban diagnosis (Liu et al. 2019, 2021b; Deng et al. 2022a, b; Wu et al. 2021; Jamonnak et al. 2022) and decision-making (Takenouchi and Choh 2021; Liu et al. 2017; Li et al. 2020; Weng et al. 2018). In essence, tour planning is an urban decision-making problem. Many visual analytics systems have been developed for challenging decision making in various urban domains. SmartAdP (Liu et al. 2017) helped visually select billboards with higher exposure potential in cities. ReACH (Weng et al. 2018) assisted citizens in finding ideal homes based on their routines and corresponding reachability constraints. WarehouseVis (Li et al. 2020) empowered experts in finding an idea warehouse to satisfy the business demands from a certain area. All three systems leveraged GPS trajectory data to estimate the situation of cities. SRVis (Weng et al. 2019) is a general method for tightly integrating spatial context in location selection. The methods above only accommodate point-based decision making but cannot be applied to our line-based tour planning scenario. More recently, some line-based approaches have also been proposed. ShuttleVis (Liu et al. 2020) helped authorities schedule customized shuttle buses to satisfy dynamic commute demands. BNVA (Weng et al. 2021) allowed experts to optimize bus networks based on the check-ins/outs of passengers. However, they are highly tailored to traffic data and cannot be applied to tourism scenarios. To this end, we design a set of new visualizations and interactions for crowdsourcing tour-based tour planning.

3 Requirement and data analysis

This section presents the background of our study, describes the relevant concepts and datasets, and summarizes user requirements that guide the design of our proposed tool.

3.1 Overview

cT^3 is an interactive tool proposed for tourists who plan to travel in a city. We refer to these people as our target users. Tour planning is a process of determining an ideal attraction sequence (tour) for each day during traveling. We realize the limitations of existing methods in terms of reliability (i.e., end-to-end) and efficiency (step-by-step), as we have mentioned in the first section. Thus, we want to empower users to plan their ideal tours efficiently and reliably. cT^3 is a tour-based tool not an end-to-end nor step-by-step tool. First, cT^3 build upon the popular attraction sequences that are extracted from crowdsourcing tour data and reflect previous tourists' experiences. Exploratory tour planning based on previous tourists' experiences enables a reliable planning process. Second, cT^3 allows users to directly manipulate multiple attractions at once with flexible interactions, which underpins efficient planning.

3.2 Concepts

We characterize the concepts involved in our tour planning tool based on the tour data as follows.

- An *attraction* l (Fig. 2A) is essentially a location. Each attraction is associated with the following attributes: latitude, longitude, attraction type (introduced below), comments, and photographs taken there. Note that our study introduces a special abstract event called "Dining." "Dining" indicates that tourists dine for rest without particular preference and geo-location.
- Eight *attraction types* are used Shi et al. (2015): geographical landscape (GL), water scenery (WS), biological landscape (BL), sky and climate landscape (SCL), ruins and monuments (RM), buildings and facilities (BF), tourist commodities (TC), and human activities (HA). These types follow the classification of tourism resources in China, where our datasets are collected.
- A *tour* T (Fig. 2B, C) comprises a set of ordered attractions $\{l\}$. Tourists will plan a tour for each day before traveling and may voluntarily share their tours on tourism web forums for reference by other tourists.
- A *subtour* T_s (Fig. 2D) is a contiguous part of a tour and has the same data structure as a tour. Intuitively, we say a tour T supports T_s if T_s can be derived from T by deleting some or no locations from the beginning and end of T. The reason we introduce this concept is that users do not follow a shared tour strictly but are more likely to follow part of it. It allows users to efficiently plan multiple orderly attractions at once. Moreover, a *subtour* (or *tour*) characterizes tourists' experiences in a structural way.
- An *itinerary* I (Fig. 2E) comprises the following two properties. (1) Several tours $\{T\}$ associated with the *city* on different days. An itinerary may involve more than one city and last several days. (2) *Copy* records how many tourists copy or like this itinerary. In fact, not all users who copy the itinerary will follow it to travel, but this metric reflects whether the itinerary is worth following.
- *Travel experience* is borrowed from Nomiyama et al.'s work Nomiyama et al. (2018). It is a subjective metric for a tourist's satisfaction with traveling. For example, if a tour is distant, tourists will feel tired; the tour of going to hot springs after skiing has a good travel experience.

Our method can be applied to crowdsourcing tour data containing attributes corresponding to the above concepts.

3.3 Requirement analysis

Before designing cT^3, we organized a brainstorming session with 8 people and interviewed other 6 people one-on-one for requirement analysis. All 14 people have experience planning their tours and belong to the target users.

First, the brainstorming session is started with the following questions: *What will you do to prepare for your travel? Do you find helpful inspiration from other tourists' itineraries or travelogues? If you were given many itineraries, how would you leverage them for your planning?* The 8 participants involved were undergraduate and graduate students. Each of them spoke freely about these questions. The brainstorming session lasted over one hour. Afterward, we summarized their comments and derived initial requirements.

Second, we interviewed other 6 people to derive concrete and systematic requirements. One participant was a senior researcher, one participant's previous major was tourism management, two had several years of expertise in urban planning, and the remaining two are travel enthusiasts who travel an average of 3 times per year. Each interview lasted 15 min. Participants commented on the initial requirements. Finally, we identified their requirements in planning ideal tours using crowdsourcing tour data as follows:

R1 *Obtain an overview of attractions visually.* Questions can be asked, *what kind of attractions are mainly in this city? Natural landscape or amusement park? How are they distributed?*

R2 *Explore tours efficiently.* Users appreciate being able to explore and leverage others' experiences to plan their own tours quickly and wisely. Yet, they do not want to be overwhelmed by a large number of tours.

R3 *Filter and compare tours effectively*. During tour planning, users will identify must-see attractions and conduct filtering accordingly. Besides, they may need to compare different tours and establish which one or which part (subtour) is more suitable.

R4 *Interpret tours convincingly*. They hope to know whether or why a tour or subtour is good. For example, *how popular it is? Does it have a good travel experience? Where do tourists go after visiting this attraction?*

R5 *Construct ideal tours interactively*. After evaluating previous tours, users need to iteratively select and conveniently connect these tours or part of them to generate final ideal tours.

R6 *Modify the tour on demand*. Users hope that our tool can support personalized attraction-level modifications, e.g., inserting a dining event or their favorite attraction into an ideal tour.

3.4 Data analysis

We adopt a frequent substring mining technique to extract popular subtours from tour data. Since the frequent substring mining is a well-established technique, we only briefly describe it below. Please refer to Appendix A for more details.

3.4.1 Substring mining

The sheer volume of tour data prevents users from easily acquiring helpful knowledge that can be used in tour planning. Itineraries are essentially movement data. According to the visual analysis methodology of movement data (Andrienko et al. 2013), a pattern extraction method is required to extract valuable knowledge from overwhelming crowdsourcing tour data (**R2**). Among various algorithms for movement data (Zheng 2015), sequential pattern mining is the most suitable one in our scenario. It can summarize massive tour data with frequent occurring subsequences of attractions, i.e., popular subtours. In this way, users can quickly capture how other tourists visited attractions. Moreover, when planning a tour, users will not follow a shared tour strictly but are more likely to follow part of it. Extracted popular subtours can well accommodate such a scenario.

We decide to adopt frequent substring mining, a kind of sequential pattern mining (Han et al. 2007), to extract popular subtours from itineraries. Briefly, a substring is a contiguous part within a sequence. A sequence can derive its substring by deleting some or no element from the beginning and end of itself. A sequence S_b supports another S_a if S_a is a substring of S_b. A substring s (a subtour) is frequently occurring if the number of sequences (tours) that support it is larger than a pre-defined threshold. We employ a well-

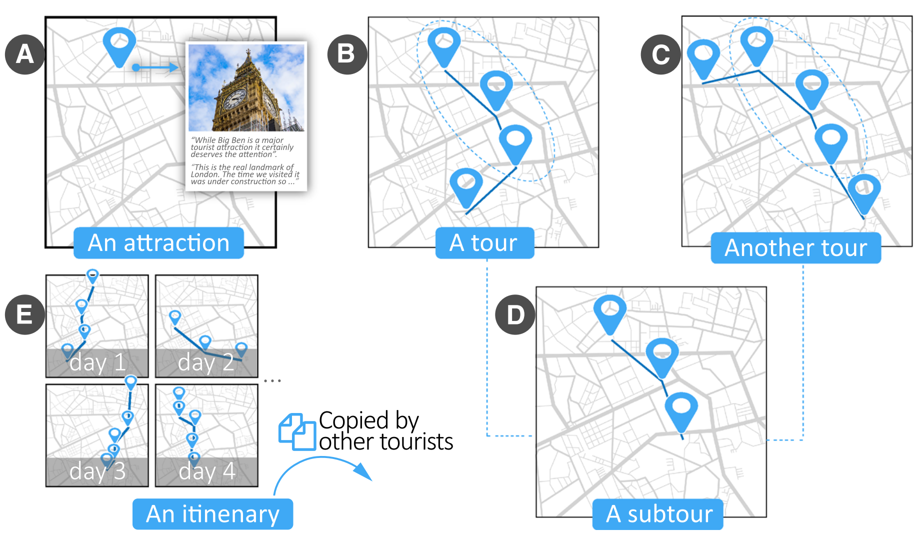

Fig. 2 Four important concepts used in this study. **A** An attraction displayed in the map and associated with some photographs and comments; **B**, **C** two illustrative tours comprising a set of ordered attractions. **D** A subtour of the both tours (**B**) and (comprises a set); **E** an itinerary comprising many tours on different days and recoding the number of times it is copied

established Apriori-based approach (Agrawal and Srikant 1994) to extract popular subtours given a collection of tours. Users can plan their own tours based on these popular subtours.

Justification. Note that the contiguity feature of frequent substrings is important in mining knowledge from tour data for tour planning. For example, an extracted subtour based on the substring mining "A → *Break* → C" indicates that tourists usually take a break (e.g., lunch) after visiting attraction A and then go to C. However, subsequence mining may extract "A → C". The break is missing, which may cause fatigue and bad travel experience (Nomiyama et al. 2018). Thus, we choose the substring mining rather than the subsequence mining.

3.4.2 *Index mechanism*

Given extracted popular subtours, we design indexes to support seamless interactions and scalable visualizations.

Parent–Children index. We say a subtour T_s^1 is another subtour T_s^2's parent (T_s^2 is T_s^1's child) if T_s^2 can be extended from T_s^1 by appending an attraction from the beginning or end of T_s^1. Parent and children subtours index each other.

Descendant index. Massive popular subtours can be extracted, which hinders effective exploration (**R2**) and analysis (**R3,R4**). So, we propose a definition of *super subtour* based on the inherent parent–children relationships and organize the popular subtours accordingly.

We define a subtour is a super subtour if it has no parent. That is, there is no popular subtour can be generated by extending a super subtour. Based on the recursive parent–children relationships, a super subtour indexes many popular subtours that belong to its descendants. A super subtour can also visually represent its indexed popular subtours because they are parts of it, which alleviates the scalability issue.

Attraction-Subtour index. For operating subtours efficiently based on attractions (**R3,R6**), we construct an attraction-subtour index. Thus, the subtours involving a given attraction can be quickly retrieved.

4 Framework and implementation

Based on the data and requirement analyses, we propose a tour planning framework as shown in Fig. 3. (1) The first step is to collect the crowdsourcing tour data, including the itineraries shared by tourists and the involved attractions (Fig. 3A). (2) Then, a frequent substring algorithm is issued to identify popular subtours from the tour data (Fig. 3B). (3) Afterward, users are allowed to learn previous tourists' experiences by visually exploring and interpreting the popular subtours (Fig. 3C). (4) Finally, users manipulate the subtours of their interests to construct and modify their ideal tours (Fig. 3D). The third step and the fourth step may alternate if users need to plan multiple tours.

We develop cT^3 following this framework. cT^3 is a web-based system for tour planning and constituted by three parts, namely data storage, backend, and frontend. First, the data storage module stores the crowdsourcing tour data. Second, the backend, implemented in Python, is used to identify popular subtours from the tour data. Third, the frontend, written in Vue.js and TypeScript, runs in modern web browsers and supports visual analysis and interactive manipulation of the identified subtours. The backend and frontend will be introduced in Sect. 5 and 6, respectively.

5 System design of cT^3

To support the gathered requirements, we design a set of visualizations and interactions, and further develop an interactive tool cT^3 (Fig. 4). cT^3 comprises a menu (A) and four views, namely map (B), tour (C), ideal-tour (D). and itinerary (E) views.

5.1 Menu

The menu (Fig. 4A) supports setting the city to visit and the tour of the day to be planned. It also shows the color scheme for the attraction types used throughout the interface. We divide the eight attraction types into two categories according to whether it is a natural scenery or historical and cultural site. For those natural scenery types, we assign colors close to green; otherwise, we assign colors close to red.

5.2 Map view

The geographic map is very common to provide the necessary spatial context for georeferenced data (Zhang et al. 2021; Zheng et al. 2021). We also adopt a map to present the attractions, subtours, and planned ideal tours.

Attraction. Each attraction is basically depicted as a circle and then colored based on the attraction type. We use the detail-on-demand mechanism to display the attractions. Initially, we encode the popularity and attraction type, with the size and color of the circle, respectively. To avoid visual clutter, those circles with low popularity are colored gray. The considerations are that (1) popularity is the most important metric for attractions, (2) while the type summarizes the content of the attraction from a high perspective. In this way, users can quickly and intuitively obtain the spatial distribution of different types of attractions in a city (**R1**). When users hover on a circle, the corresponding photographs and comments will be shown (Fig. 4b2), and the corresponding visual elements in the tour view will be highlighted.

(Sub) Tours. For simplification, we use tour to denote subtour and tour in the following, as they have the same data structure. A tour comprises ordered attractions. We visualize a tour with a graph representation (Zhao et al. 2021) where the involved attractions are directly linked with curved arrows (Fig. 4b1).

Note that drawing the tours on the map is to help users understand the tours through spatial context (**R3**), rather than allow users to explore and plan tours based on this. Attractions are fixed by their geographic coordinates. Thus, the tours on the map have various and irregular structures, making tour comparison difficult. Besides, displaying many connected arrows on the map inevitably results in visual clutter and overlapping, which prevents users from exploring, interpreting, and manipulating tours effectively.

5.3 Tour view

The tour view, the major view of the tool, is designed for visualizing numerous popular subtours and interacting with them. The tour view and the map view constitute multiple coordinated views for tour data. Compared with integrated views, such an approach affords more design space. Besides, it displays information separately to avoid information overload (Tominski et al. 2021), which is friendly to general users. Below, we describe the tour view from two aspects, namely visualizations and interactions.

5.3.1 Visualizations

The tour view adopts a uniform layout where the visual elements are neat in the vertical direction, which organizes the subtours well for efficient exploration (**R2**) and facilitates comparison (**R3**). We also encode popularity-related metrics such that users can visually evaluate and interpret them (**R4**).

In particular, scalability is the most challenging issue for visualization. To this end, we further identify super subtours from these subtours, as we described in Sect. 3.4.2. Each super subtour can visually represent their descendant subtours because they all belong to a part of the super subtour. Consequently, we reduce the number of subtours to be visualized and thereby alleviate the scalability issue.

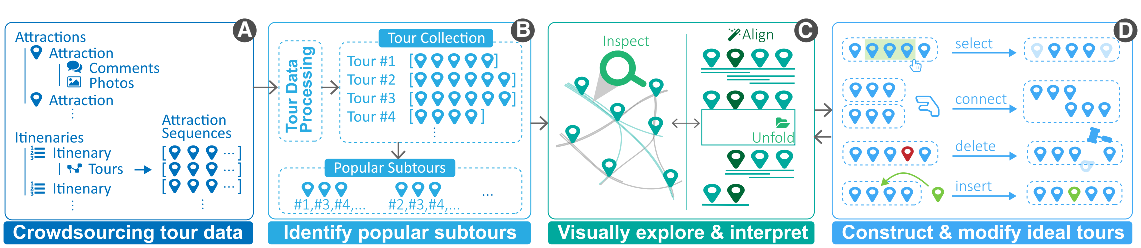

Fig. 3 Tour-Based tour planning framework. From **A** the collected crowdsourcing tour data, **B** a frequent string mining algorithm identifies numerous popular subtours. **C** Visualizations support subtour exploration and interpretation, such that users can learn previous tourists' experience. Based on these subtours and the learned knowledge, **D** users construct and modify their ideal tours

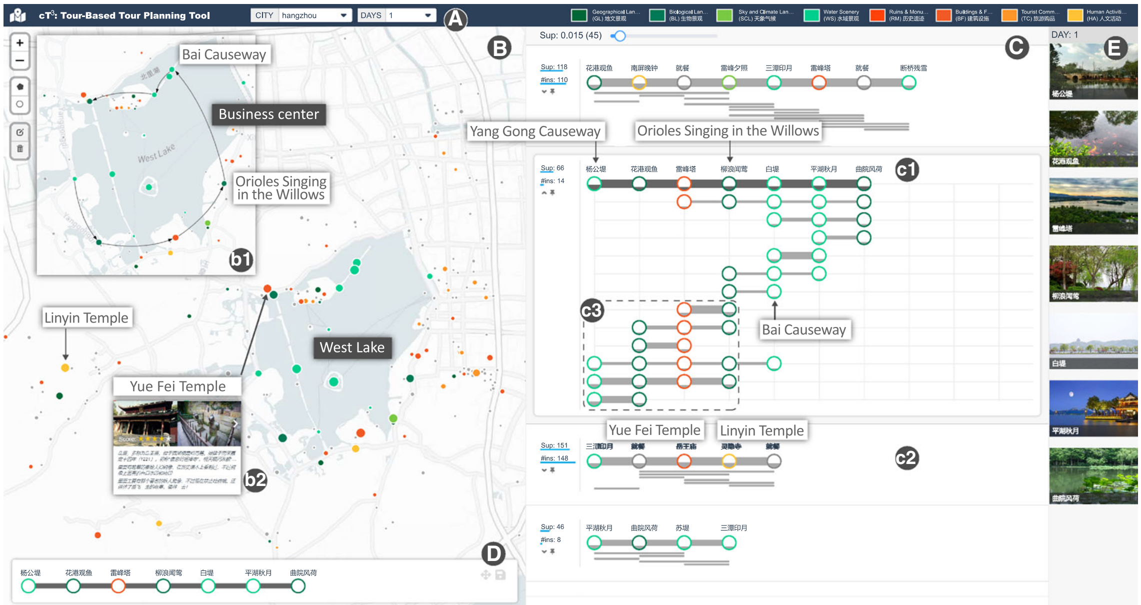

Fig. 4 The interface of cT^3. **A** The menu supports setting the travel city and the tour of the day to be planned and also shows the color legend; **B** the map view displays the attractions and tours in the spatial context; **C** the tour view visualizes the numerous popular subtours identified from the crowdsourcing tour data and implements useful interactions for tour planning; **D** the ideal-tour view presents the ideal tours selected from the tour view by users; **E** the itinerary view shows the final itinerary with planned tours

Based on the descendant relationship, we adopt a two-level hierarchy strategy. In the first level, we only visualize super subtours. In the second level, we unfold a super subtour by visualizing its descendant subtours. We describe the view from the two aspects.

Visualizing many super subtours. The tour view contains a list of tour cards (Fig. 4c1, c2) designed for super subtours. Each tour card has two states, opened and closed. Initially, the tour cards are closed (Fig. 4c2). A closed tour card only presents the super subtour as well as visual hints of its descendant subtours.

On the left side, we explicitly display two popularity metrics, namely numbers of tours and itineraries supporting this super subtour. The lengths of the blue underlines also encode the values, facilitating the comparison across different cards of super subtours.

On the right side, the super subtour is visualized as a horizontal node-link diagram. First, the nodes represent the attractions. Their names are placed above them, and their types are encoded with the colors of the node border (Fig. 6D). The attractions occurring in the super subtours have sufficient popularity, providing an attraction overview (**R1**). Furthermore, for each attraction, we calculate how many tourists who visited this attraction would follow this subtour, i.e., the popularity of the subtour divided by the popularity of the attraction. Afterward, we encode the ratio with the area of the inner part into each attraction node (Fig. 6D) because people first perceive the area (Nguyen et al. 2021). Such a ratio reflects how much the subtour is worth following. Second, the links imply the visiting order from left to right. Each link has the same length to facilitate alignment and comparison. Yet, we do not encode the popularity of the super subtour again when the card is closed. When the card is opened, i.e., in the second level, their visual channels will be utilized for further analysis, for example, comparison with its descendant subtours.

Below the super subtour are its descendant subtours. When the card is closed, there is limited space to fully visualize these subtours as node-link diagrams. So, we compact them as thin lines. The lines of the subtours are aligned with its super subtour based on their shared attractions. These compact lines serve as visual hints.

Unfolding a super subtour. The opened card (Fig. 4c1) has more space such that the super subtour can be unfolded. Particularly, the descendant subtours are visualized as node-link diagrams like the super subtour. Moreover, we encode popularity for each subtour, including the super and descendant subtours, with the thickness of the links. The popularity metrics of the subtours have been normalized within a tour card. So, users can easily compare them.

However, thick lines will be biased towards those short subtours, for example, in Fig. 5A and b1, because the longer the subtour, the more difficult it is to completely appear in tours. We notice that a tour that supports the child subtour must support its parent subtour. For example, in Fig. 5b1, Tour #1-3 support both the child and parent subtours, which leads to duplication. Thus, we refine the support of the subtour by removing the supporting tours that have been in its child subtour. For example, after refinement, the second subtour in Fig. 5b2 is only supported by Tour #4.

Refined supports produce readable visualizations. It addresses the popularity bias between short and long subtours and makes the patterns noticeable. For example, the subtour of Fig. 5c1 is obvious, indicating that many people only visited the two middle attractions; the subtour of Fig. 5b3 is also attractive, showing that people tend to visit the three attractions consecutively.

5.3.2 *Interactions*

We implement useful interactions for exploration and planning. As we present the subtours with sequence visualizations, we refer to the design space of sequence interactions proposed by Guo et al. (2021a) when designing interactions.

Hovering. Users can hover on the attraction nodes. The hovered attraction will be highlighted across different views in the interface. The corresponding photographs and comments also will be shown (Fig. 4b2). In addition, users can hover on the links of a subtour, and then the map view will displays the subtour (e.g., Fig. 4b1). Thereby, users can obtain the spatial context of subtours for decision making.

Expanding & Collapsing. Users can click the left side of a tour card to open (expand) it (Fig. 4c1). Users can collapse it by clicking again or expanding another one. This interaction helps to explore numerous subtours in the two-level hierarchy (**R2**).

Filtering. Users can filter super subtours (**R3**) based on attractions by clicking the “filter” icons above the attractions (Fig. 6a3). Afterward, the super subtours that do not involve the attraction will be filtered out. In this way, users can specify some particular attractions, for example, must-see attractions. Such a filtering interaction can also be triggered by clicking circles of attractions in the map view.

Alignment. After filtering, users can further align the subtours by clicking the “align” icon (Fig. 6a3). In Fig. 6, the five super subtours are aligned based on “Three Pools...”. Alignment unveils the positions of the reference attraction across different subtours (**R3**), enabling users to know which subtours can be connected (**R5**).

Mining. Users can change the support threshold via the slider on the top of the tour view (Fig. 4C). If the user prefer more popular subtours, s/he can raise the threshold.

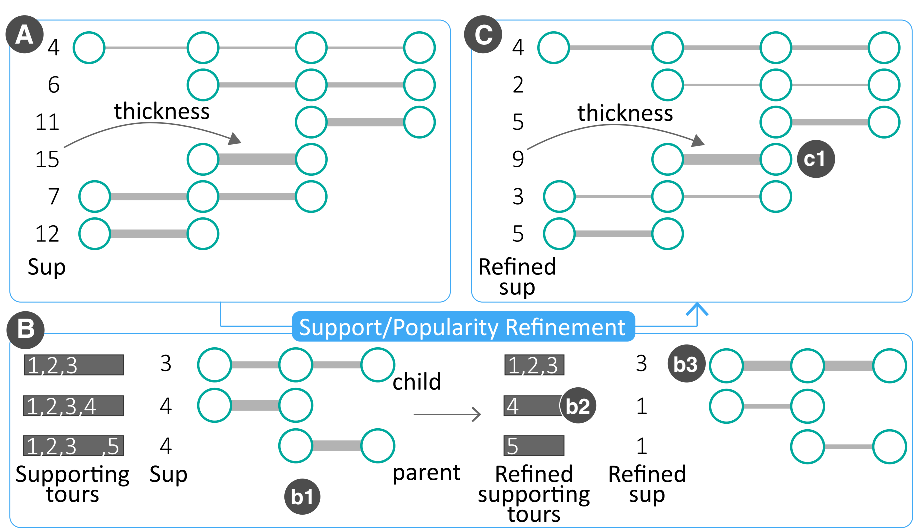

Fig. 5 Two examples illustrate support refinement for unfolding a super subtour. Particularly, **A**–**C** constitute the first example showing a tour card before and after support refinement. **B** is a simple but more concrete example showing the details of refinement

The prior interactions (Guo et al. 2021a) cannot fully satisfy our collected requirements because we are the first to attempt to utilize given sequences (i.e., subtours) to construct new sequences (i.e., final tours). We propose and implement the following new interactions.

Brushing & Selecting. Users can select the subtour as his ideal tour by brushing a range on the subtour (**R5**). The brushed subtour will be highlighted with darker gray (e.g., see Figs. 4c1 and 7b1) and visualized in the ideal-tour view (Fig. 4D).

Updating. Whenever finishing planning for a day, the attraction in the planned tour will be stored as the visited attractions. When starting a new day's planning, the tour view will be automatically updated. The updated subtours will not include those visited attractions because users usually do not visit the same attraction twice. In this way, we reduce the workload of users in evaluating and selecting the subtours (**R5**).

5.3.3 Justification

A tour can be cast as an event sequence. Each attraction in a tour can be viewed as a visiting event. Therefore, the tour view follows the design space of visual analytics techniques for event sequence data (Guo et al. 2021b). Guo et al. (2021b) summarized five types of representations for event sequences, namely chart-, timeline- (Wang et al. 2022), hierarchy- (Wongsuphasawat et al. 2011), sankey- (Wang et al. 2021b; Wu et al. 2022), and matrix-based (Wu et al. 2018) visualizations. We choose the timeline-based since it is the most intuitive visualizations. Guo et al. (2021b) also concluded seven types of interactions, filter/query, editing, segmentation, alignment, scaling, emphasis, and aggregation. To satisfy the requirements, we implement filter, query, emphasis, and alignment, which enable users to interact with the tours flexibly.

5.4 Ideal-tour view

The ideal-tour view presents the ideal subtours selected from the tour view. These ideal subtours are visualized as node-link diagrams like that in the tour view but without visual encodings. We also propose new interactions for deriving final tours in this view.

Merging & splitting. A final tour may be composed of multiple ideal subtours. Thus, we allow users to merge subtours in this view (**R5**). Two subtours can be merged if (1) they have no common attractions or (2) overlap one or more attractions at their end or beginning. For example, the ideal tour in Fig. 6C is composed of two ideal subtours in Fig. 6A, B. If two ideal subtours share the same attractions, we assume that users want to merge them because tourists usually do not visit the same attraction twice. Hence, the tool will automatically merge the subtours by default. For the two subtours with no common attractions, users can merge them by dragging one to the beginning or end of the other. Accordingly, splitting is also supported.

Modifying. Users can modify the subtours by adding or deleting attractions (**R6**) (Fig. 7D). Specifically, users can click the links of ideal subtours (Fig. 7d1), and then, a window will pop up (Fig. 7d2). Via the window, users can conduct any attraction-level modification, e.g., adding a dining event after the link (Fig. 7d3).

Saving. Users can click the "save" icon to save the ideal tour as the final tour in the itinerary view for the current day.

5.5 Itinerary View

The itinerary view (Fig. 4E) presents the final planned tours day by day. For a final tour on a certain day, we display the involved attractions from top to bottom. Here, each attraction is represented by its name and a photo taken there.

6 Evaluation with usage scenarios

We evaluate the effectiveness and usability of cT^3 with two usage scenarios on Hangzhou and Guangzhou, two famous tourism cities in China. To confirm the correctness of the obtained itineraries, we have interviewed 2 local residents for each city, and they reported that these itineraries were reasonable and well designed. We do not conduct a comparative evaluation. First, comparison under the same results is difficult because users will have different plans in different tools. Second, most of the existing tour planning tools are not open source. We believe the comparison discussed in the first two sections is sufficient to distinguish our approach from the existing ones.

6.1 Dataset preparation and pre-processing

The datasets on Hangzhou and Guangzhou are collected from one of the most popular tourism forums in China. Tourists upload their itineraries in a structured way, e.g., “Day1: A - B - C.” Thus, we can extract structural attraction sequences from the website. Before applying the mining algorithm, we manually unify the names of the same attractions and replace visiting those less known restaurants with “Dining.”

Consequently, the Hangzhou dataset comprises 114 attractions, 1,721 itineraries, and 3,584 tours; the Guangzhou dataset comprises 47 attractions, 838 itineraries, and 1,431 tours. We apply the substring mining on these two datasets to extract popular subtours. The detailed properties comprised by the data and mining results are described in Sec. 3.2. By default, the thresholds for frequent occurrence are 45 and 30, respectively. The thresholds can be interactively increased via the system interface. The Apriori-based approach (see Appendix A) is implemented with Python3. Given the thresholds, the runtimes are 11 s and 2.2 s on a desktop running Ubuntu 20.04 with Intel Core i7 3.70GHz CPU, and 16 GB RAM. To deal with larger datasets, more efficient approaches (Ji and Bailey 2007; Lee and Raedt 2004) can be adopted.

6.2 Usage scenario on Hangzhou

Tom is a travel enthusiast and likes to travel during his vacation. He has never been to Hangzhou. In this section, we follow Tom to illustrate how our tool assists a user in understanding the crowdsourcing tour data and then planning his ideal tours in Hangzhou.

Tom first navigated the map view (Fig. 4B). He noticed that most of those circles with large sizes were green, while only some of them were red. Tom also observed that the green circles were located around a big lake named West Lake. The observations show that tourists traveling in Hangzhou will go to the attractions of natural scenery, especially water scenery. He browsed the photographs and comments of the remaining attractions by hovering on the red circles (Fig. 4b2). Afterward, he knew that there were some buildings related to sacrifices (e.g., Yue Fei Temple) and religious activities (e.g., Lingyin Temple). Yue Fei Temple and Lingyin Temple were also observed in the popular subtours (Fig. 4c2), consistent with the map view. Finally, Tom learned that Hangzhou is a tourist city dominated by natural scenery, especially water scenery, as well as some religious and cultural attractions.

Day 1: Plan a long tour around West Lake. After obtaining an overview of the attractions in Hangzhou, Tom hopes to travel West Lake and visit the surrounding attractions in the afternoon of the first day. In the tour view, the second-longest popular super subtour (the first top subtour in Fig. 4c1) attracted his attention. We denote it as **Subtour #1** in the following for better description. Along with Subtour #1, Tom can travel around West Lake exactly (Fig. 4b1) and visit seven popular attractions. Tom thought that Subtour #1 was a good option but worried it was too long and tiring because he only had one afternoon.

He unfolded the tour card of Subtour #1 to analyze how other tourists travel along with it in detail (Fig. 4c1). The opened tour card displayed other popular subtours contained in Subtour #1. These shorter contained subtours may lead to a more relaxed and better travel experience. From the tour card, Tom noticed that those thicker node-link diagrams mainly appear in the front part of Subtour #1. This indicated that tourists prefer to travel on the front part, i.e., from “Yang Gong Causeway” to “Orioles Singing in the Willows” (Fig. 4c3). From the map view, Tom also identified between the “Orioles Singing in the Willows” and “Bai Causeway” is business center where many small popular cafes located. Therefore, Tom decided to travel around West Lake along with Subtour #1. If he gets tired at the attraction “Orioles Singing in the Willows” in practice, he will go to the business center for a break.

Finally, Tom brushed Subtour #1 and saved it. He finished the tour planning of the first day and started to plan for the next day.

Day 2: Connect two tours, and travel West Lake and temples. The theme of the second day’s travel is the history and culture of Hangzhou. After Tom selected day “2”, the popular subtours containing the attractions visited on the first day will be filtered out. After filtering, the longest popular subtour (Fig. 6A) happen to involved two of the most popular religious and cultural attractions, Yue Fei Temple and Lingyin Temple (Fig. 6a1). We denote it as **Subtour #2**. Tom noticed that Subtour #2 involves two “Dining” events (Fig. 6a2) and the aforementioned two temples are between them. Thus, Tom knew that tourists usually go to the two attractions in the afternoon. Tom gladly chose Subtour #2 by brushing it.

The subtour has only one attraction in the morning, which makes Tom a bit dissatisfied. To extend the tour in the beginning, Tom filtered out the super subtours that do not involve the first attraction (i.e., “Three Pools Mirroring the Moon”). Then, he further aligned the remaining popular subtours based on the first attraction. In this way, three subtours that can be connected to the beginning of Subtour #2 were clearly revealed. Among them, Tom selected the subtour from “Su Causeway” to “Three Pools Mirroring the Moon” (denoted as Subtour #3) (Fig. 6B). It was because the inner parts of the nodes in Subtour #3 had high heights (Fig. 6D). This showed that many tourists travel along with Subtour #3, which means Subtour #3 is likely to result in a good travel experience. In fact, at the end of “Su Causeway”, there is a cruise ship dock where tourists can take a cruise. It is very convenient to go to the attraction “Three Pools Mirroring the Moon”.

After Tom brushed Subtour #3, Subtour #2 and #3 were automatically merged, constituting an ideal tour (Fig. 6C). Finally, Tom successfully and effectively planned a two-day Hangzhou tour. Tom was able to enjoy the different types of attractions in Hangzhou.

6.3 Usage Scenario on Guangzhou

John is afraid of tiredness and likes to travel comfortably and loosely. He want to plan a two-day travel in Guangzhou. In this section, we follow John to further demonstrate how our tool helps users to plan their tours with individual preferences. This usage scenario also shows that our proposed framework and tool can be applied to various cities.

John preferred those popular subtours and didn’t want to visit too many attractions in one day. Therefore, he raised the support threshold and obtained more popular and shorter subtours. Then, John observed the overview of the attractions in Guangzhou. In both the map and tour views, most of the popular attractions in the map were red (Fig. 7A). By simply browsing these attractions, John learned that the major tourist attractions of Guangzhou belong to architectural landscapes (e.g., museum) and human activities (e.g., special restaurants) with local characteristics.

Day 1: Modify a tour and obtain a relaxed one.

After knowing the basic information of Guangzhou, John started to analyzed the tour data and then plan his tours in the tour view. John was very happy to identify the subtour with the largest support (Fig. 7B). We denoted this subtour as **Subtour #4**. After John hovered Subtour #4, the spatial context of Subtour #4 is displayed in the map view (Fig. 7C). It spanned only a few blocks and the total spatial distance was not far. However, John was a little dissatisfied, because 1) only walking is suitable in this small area, but 2) walking

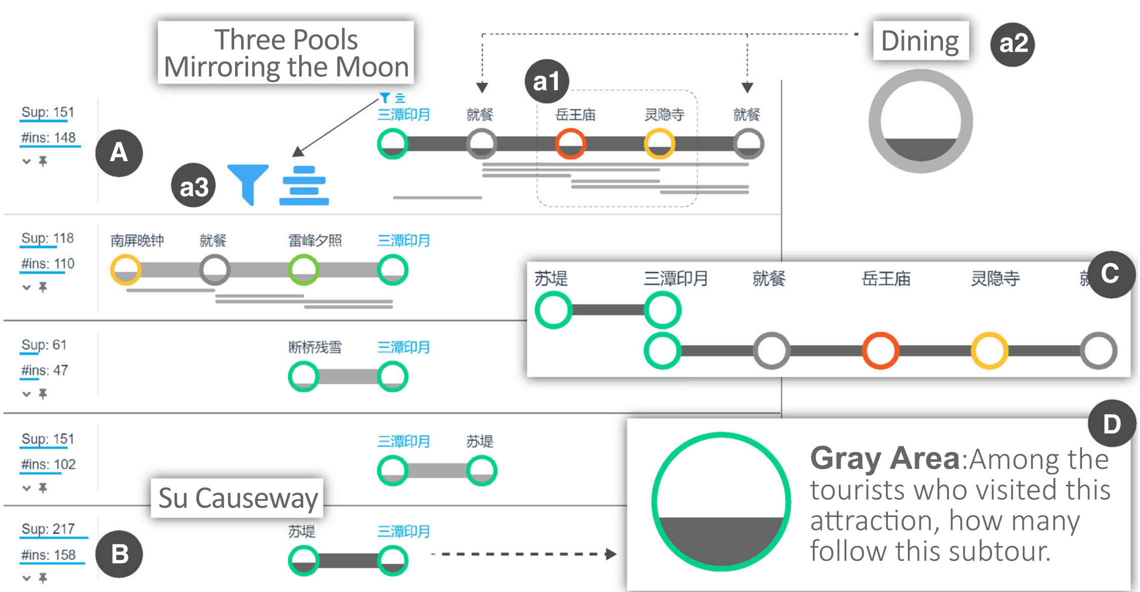

Fig. 6 Planning the tour for day 2 in Hangzhou. The left side is several subtours aligned by “Three Pools Mirroring the Moon”. **C** The final tour connected by two subtours **A**, **B** in the left side. **D** The attraction node shows that about one-third of tourists visit “Su Causeway” before visiting “Three Pools Mirroring the Moon.”

across these attractions may make him tried. Thus, he wanted to discard the first attraction “Chen Clan Academy” because the distance between it and the second attraction occupied half of the total distance.

ToIn order to see the other tourist’s choices, John unfolded the tour card of Subtour #4. The visualization supported his idea. Specifically, the contained subtours that involved the attractions at the end of Subtour #4 were almost as thick as Subtour #4 (Fig. 7b1). This indicated that many tourists only followed the back part of Subtour #4, for example, from “Yuexiu Park” to “Beijing Road”. Given that, John brush Subtour #4 from “Yuexiu Park” to “Beijing Road” and this subtour was added into the ideal-tour view.

Afterward, John planned to visit “Yuexiu Park” in the morning and the remainder in the afternoon. Thus, John required dining after visiting “Yuexiu Park” and hence needed to modify the tour (Fig. 7D). He clicked the link after “Yuexiu Park” (Fig. 7d1) and inserted a dining-event (Fig. 7d2). Finally, an ideal relaxed tour for the first day was successfully derived and saved (Fig. 7d3).

Day 2: Directly select a tour with good travel experience.

John started to plan for the second day. After the tour view was updated, only a few subtour were left. John noticed the subtour (denoted as **Subtour #5**) involving four attractions and two dining events (Fig. 7E). Subtour #5 implied visiting two attractions in the morning and other two in the afternoon. John thought it acceptable.

He directly selected and saved it as his second-day’s tour after analyzing the travel experience from the following two aspects. (1) Popularity. The tour card of Subtour #5 unveiled two thickest subtours involving the attractions to be visited in the morning and afternoon, respectively (Fig. 7e1 and e2). The thicknesses showed that many tourists visited the previous attraction and then visited the next attraction and hence implied good travel experience. (2) Spatial distance. Both the spatial distances between the first two attractions and the last two were close (Fig. 7F). John, who is afraid of tiredness, highly appreciated this. Being close may also be the main reason for good experience for tourists.

7 Discussion

This study has two **implications**. First, it is the first step towards tour planning using crowdsourcing tour data. We propose a tour-based tool planning framework and accordingly implement a novel interactive tool. Such a tour-based approach features a balance between efficiency (end-to-end) and reliability (step-by-step). The subtours identified by the pattern extraction model authorize users to select multiple attractions at once and thereby ensure efficiency, while the proposed interactive visualizations provide a transparent and confident environment for tour planning, which improves reliability. Second, we prove the value of crowdsourcing tour data, which brings enlightenment to those tourism forum or platform managers. Currently, on many forums, itineraries are only shared in an unstructured text form. Extracting structured and ready-to-use itinerary data from such text via NLP techniques remains a challenge. To collect high-quality data, managers can intentionally encourage users to share their itineraries and provide useful user interfaces for collecting data with more dimensions in a structured way. Afterward, higher-quality data can be used for better tour planning, which can improve the functionality of the platforms.

The gained **insights** during proposing this approach can be concluded as follows. First, the data on the Web cannot be directly used. Filtering is needed. For example, not all visitors to forum websites share their itineraries and comments sincerely and carefully. In this study, we filter out those short tours (less than 3 attractions) and short comments (less than 10 words). Second, no matter how powerful models are, the human intelligence should be retained in a practical and useful system. The popularity adopted in this study

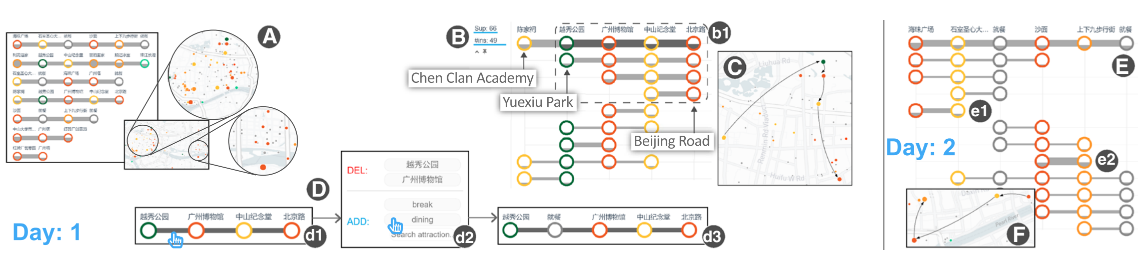

Fig. 7 Tour planning in Guangzhou. **A** Overview of attractions. **B** A tour card and **C** the spatial information of the super subtour with the largest support. **D** Personalized modification of an ideal subtour. **E** The tour card and **F** the spatial information of a subtour with good travel experience

can be easily understood and accepted by users. Besides, to further ensure practicality, we adopt multiple coordinated views and level-of-detail mechanism to support exploration and interpretation.

Three **limitations** are observed in this study. First, the interface of cT^3 adopts multiple coordinated views, which prevents users from exploring and utilizing the subtours well with geographic context to some degree. Nevertheless, we believe it is the best solution to enable general users to successfully complete the complex requirements. Weng et al. (2019) propose SRVis, a novel visualization for integrating spatial context for point-based decision making. Yet, SRVis cannot accommodate our line-based scenario. We think this problem of integrating spatial context tightly for line-based decision making is valuable and plan to study it in the future. Second, the functionality of the tool needs to be improved, such as booking hotels and tickets, checking the opening hours of attractions. However, we believe it is beyond the research scope of this study, which mainly focuses on tour-based tour planning. In the future, we will be committed to cooperating with commercial companies to promote and deploy this new tourism planning paradigm. Third, the system can be further validated by user studies. In the future, we plan to implement prototypes of previous approaches for comparison. Users could be required to plan their tours freely or complete some tasks using different approaches. Besides, user feedback could be collected with post-study questionnaires. Such that, we can understand our approach's weaknesses and advantages compared with previous approaches and thereby further improve it.

8 Conclusion

We study how to make better use of tour data on the web. To enable users to leverage the previous tourists' experience for tour planning, we propose a concept of subtour and employ a frequent substring mining technique to identify popular subtours from the large-scale tour data. Based on the popular subtours, we compile user requirements and propose a tour-based tour planning framework, which guides us to develop a novel interactive tool cT^3. To our knowledge, we are the first to study empowering users to plan their ideal tours efficiently and reliably with prior tourists' experiences.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12650-022-00884-1.

Acknowledgements We thank all reviewers for their constructive comments. We also thank Huachang Yu for his contribution in data collection. The work was supported by NSFC (62072400) and the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU). This work was also partially funded by the Zhejiang Lab (2021KE0AC02).

References

Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. In: Proceedings of the international conference on very large data bases, pp 487–499

Andrienko GL, Andrienko NV, Bak P, Keim DA, Wrobel S (2013) Visual analytics of movement. Springer, Berlin

Andrienko NV, Andrienko GL, Miksch S, Schumann H, Wrobel S (2021) A theoretical model for pattern discovery in visual analytics. Vis Inform 5(1):23–42

Bai J, Zhang H, Qu D, Lv C, Shao W (2021) FGVis: visual analytics of human mobility patterns and urban areas based on f-glove. J Vis 24(6):1319–1335

Brennan S, Meier R (2007) STIS: smart travel planning across multiple modes of transportation. In: Proceedings of IEEE intelligent transportation systems conference, pp 666–671

Chen S, Yuan X, Wang Z, Guo C, Liang J, Wang Z, Zhang XL, Zhang J (2016) Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data. IEEE Trans Vis Comput Gr 22(1):270–279

Claudio P, Yoon S (2014) Metro transit-centric visualization for city tour planning. Comput Graph Forum 33(3):271–280

Contractor D, Goel S, Mausam, Singla P (2021) Joint Spatio-textual reasoning for answering tourism questions. In: Proceedings of the world wide web conference, ACM/IW3C2, pp 1978–1989

Dadoun A, Troncy R, Ratier O, Petitti R (2019) Location embeddings for next trip recommendation. In: Companion of proceedings of the World Wide Web conference, ACM, pp 896–903

Deng Z, Weng D, Chen J, Liu R, Wang Z, Bao J, Zheng Y, Wu Y (2020) AirVis: visual analytics of air pollution propagation. IEEE Trans Vis Comput Graph 26(1):800–810

Deng Z, Weng D, Liang Y, Bao J, Zheng Y, Schreck T, Xu M, Wu Y (2022) Visual cascade analytics of large-scale spatiotemporal data. IEEE Trans on Vis Comput Graph 28(6):2486–2499

Deng Z, Weng D, Xie X, Bao J, Zheng Y, Xu M, Chen W, Wu Y (2022) Compass: towards better causal analysis of urban time series. IEEE Trans Vis Comput Graph 28(1):1051–1061

Deng Z, Weng D, Liu S, Tian Y, Xu M, Wu Y (2023) A survey of urban visual analytics: advances and future directions. Comput Vis Media. https://doi.org/10.1007/s41095-022-0275-7
Dunstall S, Horn MET, Kilby P, Krishnamoorthy M, Owens B, Sier D, Thiébaux S (2003) An automated itinerary planning system for holiday travel. Inf Technol Tour 6(3):195–210
Google (2022) Google travel. https://www.google.com/travel/. Accessed 26 Apr 2022
Guo Y, Guo S, Jin Z, Kaul S, Gotz D, Cao N (2021) Survey on visual analysis of event sequence data. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2021.3100413
Guo Y, Guo S, Jin Z, Kaul S, Gotz D, Cao N (2021) A survey on visual analysis of event sequence data. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2021.3100413
Han J, Cheng H, Xin D, Yan X (2007) Frequent pattern mining: current status and future directions. Data Min Knowl Discov 15(1):55–86
Herzog D, Sikander S, Wörndl W (2019) Integrating route attractiveness attributes into tourist trip recommendations. In: Companion of proceedings of the world wide web conference, ACM, pp 96–101
Hu F, Li Z, Yang C, Jiang Y (2019) A graph-based approach to detecting tourist movement patterns using social media data. Cartogr Geogr Inf Sci 46(4):368–382
Inspirock (2022) Trip Planner: plan & manage your vacation itinerary on Inspirock. https://www.inspirock.com/. Accessed 26 Ap 2022
Jamonnak S, Zhao Y, Huang X, Amiruzzaman M (2022) Geo-context aware study of vision-based autonomous driving models and spatial video data. IEEE Trans Vis Comput Graph 28(1):1019–1029
Ji X, Bailey J (2007) An efficient technique for mining approximately frequent substring patterns. In: Workshops proceedings of ICDM, pp 325–330
Kádár B, Gede M (2013) Where do tourists go? Visualizing and analysing the spatial distribution of geotagged photography. Cartogr Int J Geogr Inf Geovis 48(2):78–88
Kinoshita Y, Yokokishizawa H (2008) A tour route planning support system with consideration of the preferences of group members. In: Proceedings of the IEEE international conference on systems, man and cybernetics, pp 150–155
Klein K, Jaeger S, Melzheimer J, Wachter B, Hofer H, Baltabayev A, Schreiber F (2021) Visual analytics of sensor movement data for cheetah behaviour analysis. J Vis 24(4):807–825
Kurashima T, Iwata T, Irie G, Fujimura K (2010) Travel route recommendation using geotags in photo sharing sites. Proc CIKM 2010:579–588
Kurata Y, Hara T (2014) CT-Planner4: toward a more user-friendly interactive day-tour planner. In: Proceedings of international conference on information and communication technologies, Springer, pp 73–86
Lee SD, Raedt LD (2004) An efficient algorithm for mining string databases under constraints. In: Proceedings of international workshop on knowledge discovery in inductive databases, vol 3377, pp 108–129
Lee JY, Tsou M (2018) Mapping spatiotemporal tourist behaviors and hotspots through location-based photo-sharing service (flickr) data. In: Krisp JM (ed) Progress in location based services. Springer, Berlin, pp 315–334
Li Q, Liu QQ, Tang CF, Li ZW, Wei SC, Peng XR, Zheng MH, Chen TJ, Yang Q (2020) Warehouse Vis: a visual analytics approach to facilitating warehouse location selection for business districts. Comput Graph Forum 39(3):483–495
Lim KH, Wang X, Chan J, Karunasekera S, Leckie C, Chen Y, Tan CL, Gao FQ, Wee TK (2016) PersTour: A personalized tour recommendation and planning system. In: Late-breaking Results, demos, doctoral consortium, workshops proceedings and creative track of the ACM conference on hypertext and social media, CEUR workshop proceedings, vol 1628
Lim KH, Chan J, Karunasekera S, Leckie C (2019) Tour recommendation and trip planning using location-based social media: a survey. Knowl Inf Syst 60(3):1247–1275
Liu QQ, Li Q, Tang CF, Lin H, Ma X, Chen T (2020) A visual analytics approach to scheduling customized shuttle buses via perceiving passengers' travel demands. In: Proceedings of IEEE visualization conference, pp 76–80
Liu D, Weng D, Li Y, Bao J, Zheng Y, Qu H, Wu Y (2017) SmartAdP: visual analytics of large-scale taxi trajectories for selecting billboard locations. IEEE Trans Vis Comput Graph 23(1):1–10
Liu D, Xu P, Ren L (2019) TPFlow: progressive partition and multidimensional pattern extraction for large-scale Spatio-temporal data analysis. IEEE Trans Vis Comput Graph 25(1):1–11
Liu H, Chen X, Wang Y, Zhang B, Chen Y, Zhao Y, Zhou F (2021) Visualization and visual analysis of vessel trajectory data: a survey. Vis Inf 5(4):1–10
Liu L, Zhang H, Liu J, Liu S, Chen W, Man J (2021) Visual exploration of urban functional zones based on augmented nonnegative tensor factorization. J Vis 24(2):331–347
Liu S, Weng D, Tian Y, Deng Z, Xu H, Zhu X, Yin H, Zhan X, Wu Y (2023) ECoalVis: visual analysis of control strategies in coal-fired power plants. IEEE Trans Vis Comput Graph 29(1) (to appear)
Nguyen VT, Jung K, Gupta V (2021) Examining data visualization pitfalls in scientific publications. Vis Comput Ind Biomed Art 4(1):27
Nomiyama M, Takeuchi T, Onimaru H, Tanikawa T, Narumi T, Hirose M (2018) Xnavi: travel planning system based on experience flows. ACM Interact Mob Wearable Ubiquitous Technol 2(1):1–25
Sharda N, Ponnada M (2008) Tourism blog visualizer for better tour planning. J Vacat Market 14(2):157–167
Shi L, Zhao H, Li Y, Ma H, Yang S, Wang H (2015) Evaluation of Shangri-la county's tourism resources and ecotourism carrying capacity. Int J Sustain Dev World Ecol 22(2):103–109
Silamai N, Khamchuen N, Phithakkitnukoon S (2017) TripRec: trip plan recommendation system that enhances hotel services. In: Adjunct proceedings of the ACM international joint conference on pervasive and ubiquitous computing and proceedings of the ACM international symposium on wearable computers, ACM, pp 412–420
Takenouchi K, Choh I (2021) Development of a support system for creating disaster prevention maps focusing on road networks and hazardous elements. Vis Comput Ind Biomed Art 4(1):22
Taylor K, Lim KH, Chan J (2018) Travel itinerary recommendations with must-see points-of-interest. In: Companion of proceedings of the World Wide Web Conference, ACM, pp 1198–1205

Thudt A, Baur D, Huron S, Carpendale S (2016) Visual mementos: reflecting memories with personal data. IEEE Trans Vis Comput Graph 22(1):369–378

Tominski C, Andrienko GL, Andrienko NV, Bleisch S, Fabrikant SI, Mayr E, Miksch S, Pohl M, Skupin A (2021) Toward flexible visual analytics augmented through smooth display transitions. Vis Inf 5(3):28–38

Travelchime Inc (2022) Wanderlog: travel itineraries and trip planner. https://wanderlog.com/. Accessed 26 Apr 2022

Tripadvisor Inc (2022) Tripadvisor: read reviews, compare prices & book. https://www.tripadvisor.com/. Accessed 26 Apr 2022

Wang H, Ni Y, Sun L, Chen Y, Xu T, Chen X, Su W, Zhou Z (2021) Hierarchical visualization of geographical areal data with spatial attribute association. Vis Inf 5(3):82–91

Wang Y, Liang H, Shu X, Wang J, Xu K, Deng Z, Campbell CD, Chen B, Wu Y, Qu H (2021) Interactive visual exploration of longitudinal historical career mobility data. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2021.3067200

Wang Y, Peng T, Lu H, Wang H, Xie X, Qu H, Wu Y (2022) Seek for success: a visualization approach for understanding the dynamics of academic careers. IEEE Trans Vis Comput Graph 28(1):475–485

Wang Q, Yin H, Chen T, Huang Z, Wang H, Zhao Y, Hung NQV (2020) Next point-of-interest recommendation on resource-constrained mobile devices. In: Proceedings of the World Wide Web Conference, ACM/IW3C2, pp 906–916

Wei D, Li C, Shao H, Tan Z, Lin Z, Dong X, Yuan X (2021) SensorAware: visual analysis of both static and mobile sensor information. J Vis 24(3):597–613

Weng D, Chen R, Deng Z, Wu F, Chen J, Wu Y (2019) SRVis: towards better spatial integration in ranking visualization. IEEE Trans Vis Comput Graph 25(1):459–469

Weng D, Zheng C, Deng Z, Ma M, Bao J, Zheng Y, Xu M, Wu Y (2021) Towards better bus networks: a visual analytics approach. IEEE Trans Vis Comput Graph 27(2):817–827

Weng D, Zhu H, Bao J, Zheng Y, Wu Y (2018) HomeFinder revisited: finding ideal homes with reachability-centric multi-criteria decision making. In: Proceedings of the ACM CHI conference on human factors in computing systems, p 247

Wongsuphasawat K, Gómez JAG, Plaisant C, Wang TD, Taieb-Maimon M, Shneiderman B (2011) LifeFlow: visualizing an overview of event sequences. In: Proceedings of ACM CHI, pp 1747–1756

Wu Y, Lan J, Shu X, Ji C, Zhao K, Wang J, Zhang H (2018) iTTVis: interactive visualization of table tennis data. IEEE Trans Vis Comput Graph 24(1):709–718

Wu Y, Weng D, Deng Z, Bao J, Xu M, Wang Z, Zheng Y, Ding Z, Chen W (2021) Towards better detection and analysis of massive spatiotemporal co-occurrence patterns. IEEE Trans Intell Transp Syst 22(6):3387–3402

Wu J, Liu D, Guo Z, Xu Q, Wu Y (2022) TacticFlow: visual analytics of ever-changing tactics in racket sports. IEEE Trans Vis Comput Graph 28(1):835–845

Yahi A, Chassang A, Raynaud L, Duthil H, Chau DHP (2015) Aurigo: an interactive tour planner for personalized itineraries. In: Proceedings of the international conference on intelligent user interfaces, ACM, pp 275–285

Yim H, Ahn HJ, Kim JW, Park SJ (2004) Agent-based adaptive travel planning system in peak seasons. Exp Syst Appl 27(2):211–222

Zhang W, Ma Q, Pan R, Chen W (2021) Visual storytelling of song ci and the poets in the social-cultural context of song dynasty. Vis Inf 5(4):34–40

Zhao Y, Shi J, Liu J, Zhao J, Zhou F, Zhang W, Chen K, Zhao X, Zhu C, Chen W (2021) Evaluating effects of background stories on graph perception. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2021.3107297

Zheng Y (2015) Trajectory data mining: an overview. ACM Trans Intell Syst Technol 6(3):1–41

Zheng F, Wen J, Zhang X, Chen Y, Zhang X, Liu Y, Xu T, Chen X, Wang Y, Su W, Zhou Z (2021) Visual abstraction of large-scale geographical point data with credible spatial interpolation. J Vis 24(6):1303–1317