


RouteVis: Quantitative Visual Analytics of Various Factors to Understand Route Choice Preferences

C. Lv^{1,2}, H. Zhang^{†1,2} , Y. Lin^{1,2}, J. Dong^{1,2} and L. Tian^{1,2}

¹Northeast Normal University, School of Information Science and Technology, China

²Key Laboratory of Intelligent Information Processing of Jilin Universities, China

Abstract

Analyzing the preference of route choice not only facilitates the understanding of individuals' decision-making behavior, but also provides valuable information for improving traffic management strategies. As the layout of the road network, the variability of individual preferences and the spatial distribution of origins and destinations all play a role in route choice, it is a great challenge to reveal the interplay of such numerous complex factors. In this paper, we propose RouteVis, an interactive visual analytics system that enables traffic analysts to gain insight into what factors drive individuals to choose a specific route. To uncover the relationship between route choice and influencing factors, we design a quantitative analytical framework that supports analysts in conducting closed-loop analysis of various factors, i.e., data preprocessing, route identification, and the quantification of influence and contribution. Furthermore, given the multidimensional and spatio-temporal characteristics of the analysis results, we customize a set of coordinated views and visual designs to provide an intuitive presentation of the factors affecting people's travels, thus freeing analysts from tedious repetitive tasks and significantly enhancing work efficiency. Two typical usage scenarios and expert feedback on the system's functionality demonstrate that RouteVis can greatly enhance the capabilities of understanding the travel status.

CCS Concepts

• **Human-centered computing** → Visual analytics; Geospatial Data; Information visualization;

1. Introduction

Driven by the rapid advancements in urban transportation infrastructure, the accessibility between regions has been greatly improved, resulting in a diversity of route choice behavior when travelling from a specific origin to a destination. Route choice refers to the process of people choosing the most suitable route from many accessible options and analyzes the factors that affect the decision [BZA17]. Quantitatively analyzing factors that influence route choice in urban transportation is of paramount importance due to its direct implications for rational urban planning and the optimization of traffic conditions. For instance, multiple routes leading to the central business district in a certain region are experiencing congestion issues. Through quantitative analysis, it may be revealed that factors such as travel time and road capacity significantly influence commuters' route choices. Armed with this information, urban planners can enhance public transportation, or implement intelligent signal control strategies to alleviate congestion and improve overall traffic flow.

The human-centered statement preference survey [ABWL98]

supports the exploration of the factors influencing route choice from an empirical perspective. However, the collected data is accompanied by subjective cognitive bias, which leads to the difference between claimed and observed behaviors. Compared with surveys, trajectories that record travel spatio-temporal information can discover more realistic human movements, which provides unprecedented opportunities for the effective and accurate analysis of route choice behavior [LSD18]. Considerable efforts have been made to analyze the authentic factors affecting route choice with trajectory data, such as using mixed path size logit method to model taxi trajectories [TWH*20]. Since route choice behavior is inherently a multifaceted decision-making process, analyzing the respective effects of different factors on it becomes quite complex. On the one hand, the diversity of influencing factors increases the complexity of determining the extent to which they affect route choice, as there may be nonlinearities, interactions, and feedback loops between these factors. Generally, route choice is influenced by route attributes and individual characteristics (e.g., congestion on roads and departure time) [DS17]. In this case, analysts may have trouble discovering which factors are most important and which can be ignored. On the other hand, route choice behavior is closely associated with the spatio-temporal context, which means that even for the same person, the determinants of route choice

[†] Huijie Zhang (zhanghj167@nenu.edu.cn) is the corresponding author.

behavior may vary when faced with different travel contexts, such as going to work or to the station. Therefore, it is a crucial task for traffic analysts to develop a flexible and versatile analytical method.

For the purpose of providing a systematic approach to analyzing route choice, we propose a quantitative analytical framework that allows traffic analysts to gain comprehensive insight into the underlying factors behind decisions. The framework integrates diverse machine learning methods (e.g., k-Gram, Ranking SVM) to realize step-by-step in-depth analysis of trajectory data, including route identification, pattern extraction, and influencing factor analysis. A typical learning-to-rank method, Ranking SVM [CXL*06], is introduced to explore the reasons why some factors are more important than others without requiring assumptions to be made about the data in advance. Given the intuitive expression and rich interactions of visual analytics, we propose a visual analytics system with multiple coordinated views to display travel information (e.g., distribution of departure time and travel patterns) from different perspectives and to support interactive exploration and comparison of what factors influence route choice in a specific context. RouteVis is equipped with flexible interactions that allow customized travel scenarios to obtain analysis results, eliminating the need to examine extensive data tables and reducing the time spent on data analysis. The multi-factor exploration view summarizes and compares the factors that affect route choice at three levels of detail: the overall weight, contribution to individual route, and raw values. The main contributions of this paper are as follows:

- We proposed a quantitative analytical framework that identifies critical routes in massive trajectories and reveals the weights of various factors on route choice.
- We proposed RouteVis, an interactive visual analytics system that utilizes rich visual elements and interactions to intuitively understand people's route choice behavior.
- We evaluated the effectiveness of the proposed methods with two usage scenarios and expert interviews demonstrating that RouteVis improves productivity of traffic analysts.

2. Related Work

2.1. Analysis of Route Choice

Analyzing route choice not only facilitates an understanding of people's preferred travel method and route, but also enables the prediction of future traffic conditions [Pra09]. Existing methods for exploring route choice can be divided into two categories: route choice modeling and visual analysis. Wardrop [War52] made the first attempt to research route choice behavior and proposed Wardrop equilibrium theory, which laid a solid foundation for subsequent research such as stochastic user equilibrium [DS77], path size logic [BAB99], cumulative prospect theory [XZX11], and so on. Scott et al. [SLB21] and Alivand et al. [AHS15] successfully applied the path size logic model to deduce the factors affecting the choice of shared bicycle and scenic routes, respectively. Furthermore, Cho et al. [CK22] devised a mixed path size correction logit model, an improvement on path size logic, to explore the route choice behavior of different traveler groups. From another perspective, Sun et al. [SZZ*14] introduced a regression method to determine whether there is a clear quantitative relationship between

route choice and other factors. Papinski et al. [PS11] deployed an analysis tool that used ArcGIS to reveal the influence of forty variables on route decisions. Yao et al. [YB20] proposed a novel approach combining random forest and a discrete choice model to explain the route choice behavior of individuals.

With the mature development of visualization and visual analysis technology, Lu et al. [LWY15] proposed a visual analysis approach to explore route choice behavior through trajectory ranking, providing insights into the travel characteristics of taxis. This work was further extended so that multiple routes could be analyzed at a time [LLY*17]. Recently, Shin et al. [SJK*23] designed a visual analytics system with exploration, modeling, and reasoning stages to support interactive route choice modeling analysis.

Most of the state-of-the-art research (e.g. Cho et al. [CK22], Shin et al. [SJK*23], etc.) focuses on using the route utility information to infer the impact of route choice. Considering the confidentiality of data, obtaining such utility value is often challenging, thereby limiting the scalability of the aforementioned methods. Lu et al. [LWY15] attempted to explore route choice without utility, which provides inspiration for the workflow of this work. They studied travel behavior along a single route, whereas we the simultaneous analysis of multiple routes, which extends their work.

2.2. Visual Exploration of Traffic Data

Transportation is one of the most important components in smart cities [CZKZZ22] and plays a vital role in promoting the economic development and social progress. Many systematic surveys have been conducted [CGW15, ZWC*16, AAC*17, MSL*19], showing that transportation-related visual analytics studies can be grouped into two categories, namely, the visual analysis of trajectories and the perception of travel patterns. Large-scale trajectory data can reflect the status of traffic and the movement of people, thus providing a basis for high-quality urban services such as billboard location selection [LWL*17], traffic congestion management [PYSJ21], and identification of urban functional areas [ZWC*18]. In terms of the exploration of trajectories, Zhu et al. [ZCX*19] proposed a situation-aware representation approach to enable the visual analysis of human mobility. AL-Dohuki et al. [ADKZ*17] made a breakthrough in using a text search engine to manage and query taxi trajectories. Likewise, Huang et al. [HZC*20] combined natural language processing and visual analysis to effectively solve the problem of uncertain trajectory queries. In addition, some researchers have been working on data abstraction [ZMT*19] and pattern extraction of origin-destination trajectories [ZFA*16].

Since visualization has an excellent ability to gain valuable information, and it is widely used to perceive and summarize inherent patterns (e.g. interchange patterns [ZFAQ13], travel patterns [HMK*20]). With respect to the exploration of mobility patterns, Qi et al. [QHGF19] proposed a multi-step methodology which analyzes and predicts mobility patterns by integrating non-negative tensor factorization and artificial neural networks. Recently, Bai et al. [BZQ*21] proposed a visual analytics system (called FGVis) to support the understanding of the relationship between urban areas.

Regarding the pattern extraction of trajectories, the traditional partition-based and density-based methods are not appropriate

for the case of feature vectors with inconsistent lengths, emerging in the proposed quantitative analytical framework. Distinguishing from existing work, our approach adopts a unified feature representation from the perspective of road sequences, enhancing the scalability of route identification. Furthermore, the visual design and interaction within the developed visual analytics systems (e.g., FGVis [BZQ*21] and SmartAdP [LWL*17]) for traffic data provided valuable references for the customization of RouteVis.

3. System Overview

3.1. Data Description and Concepts

The trajectory dataset is downloaded from Didi Chuxing GAIA Initiative, covering longitude from $104.0478^{\circ}E$ to $104.138^{\circ}E$ and latitude from $30.658^{\circ}N$ to $30.734^{\circ}N$ from November 1 to November 30, 2016. The dataset consists of 5,137,861 records, each of which includes an order ID, origin coordinates, destination coordinates, timestamp, and location point. The data can reflect people's travel characteristics at different times of the month. We explain the relevant concepts in detail (shown in Figure 1), which can better facilitate the understanding of proposed framework.

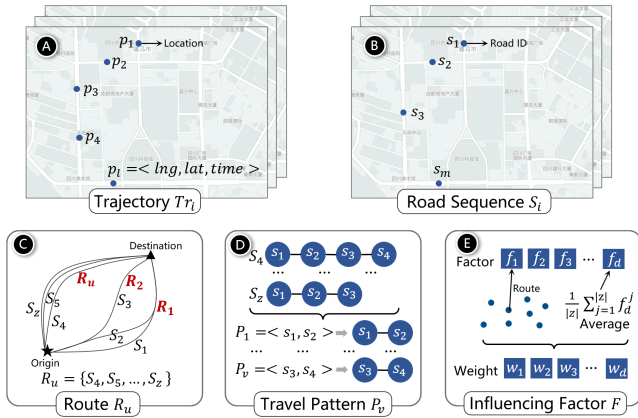


Figure 1: Illustration of concepts.

Trajectory: A trajectory consists of several locations (Figure 1(A)), denoted as $Tr_i = \{p_1, p_2, \dots, p_l\}$, where p_l is a triple with longitude, latitude and timestamp; l is the number of locations.

Road sequence: We transform the location p_l to a road ID based on its geographic coordinate, which can provide a higher level of geolocation information to understand the behavior of moving objects in spatial context. Hence, road sequence is an ordered collection of road IDs (Figure 1(B)), expressed as $S_i = \{s_1, s_2, \dots, s_m\}$, where s_m is a road ID and m represents the number of roads. Due to the short sampling time (usually 3 to 5 seconds), multiple points in the trajectory may be mapped to the same road, and we keep just one road ID.

Route: A route typically refers to a way of getting from one place to another, and is denoted as $R_u = \{s_1, s_2, \dots, s_z\}$, where z represents the number of road sequences (Figure 1(C)). To alleviate the analyst's workload and discover common influencing factors, we merge similar road sequences into a route with overlapping passing roads.

Table 1: Detailed description of influencing factors.

Object	Factor	Description
Trajectory	Departure Time	The timestamp of the trajectory departure: dawn (0:00 ~ 6:00), morning (6:00 ~ 12:00), afternoon (12:00 ~ 18:00), evening (18:00 ~ 0:00).
	Travel Cost	The duration from one place to another.
	Travel Distance	Distance between origin and destination.
Road	Road Importance	Average of road level (derived from OpenStreetMap [†]).
	Intersection Number	Number of intersections passed.
	Congestion Degree	There are four types of congestion: heavy (less than 20km/h), mild ([20,30)km/h), congested ([30,40)km/h), smooth (more than 40km/h) [11520].

Travel pattern: The travel pattern summarizes the regularity of travel and can be described by the continuous segment of the road sequence (Figure 1(D)), e.g., passing through "Li Hua Street-Shan Xi Street-Jun Ping Street". The formal representation is $P_v = \langle s_q, \dots, s_{q+\theta} \rangle$, where s_q and $s_{q+\theta}$ are the beginning and end of the road sequence and θ is the length.

Influencing factors: Influencing factors are quantitative indicators that affect route choice (Figure 1(E)), defined as $F = \{f_1, \dots, f_d\}$, where $f_d = \frac{1}{|z|} \sum_{j=1}^z f_d^j$ represents a factor, whose value is the average of all trajectories. z represents the number of road sequences. The influencing factors are summarized in Table 1, which is distilled from literature and a preliminary questionnaire conducted among the general public.

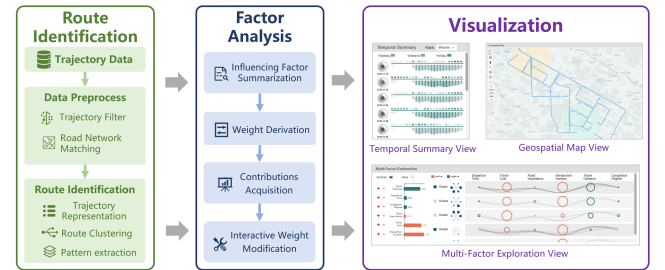


Figure 2: Pipeline of quantitative analytical framework.

3.2. The Pipeline of Analysis

We propose a quantitative analytical framework to quantitatively analyze the factors affecting route choice, which consists of

[†] <https://www.openstreetmap.org>.

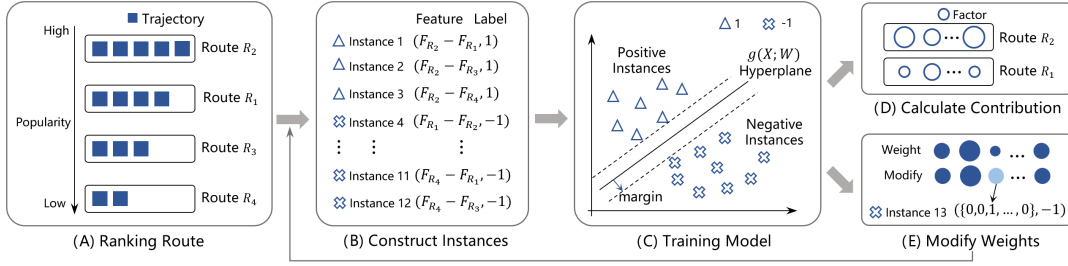


Figure 3: Workflow for quantitative analysis of influencing factors.

three modules: route identification, factor analysis and visualization (shown in Figure 2). The route identification module is primarily responsible for obtaining routes through integrated representation, clustering, and pattern mining. The factor analysis module applies Ranking SVM to derive the respective weights of factors and further utilizes simple average weighting to obtain the corresponding contributions, while allowing the custom modification of the desired weight. The visualization module comprises five coordinated multiple views to intuitively present the results of the analysis, which are implemented using Django, Vue.js and D3.js. The source code of RouteVis is available at <https://github.com/mmccc/RouteVis.git>.

4. Perceiving the Preference of Route Choice

4.1. Identification of Critical Route

Data preprocessing and representation. To ensure the accuracy and reliability of the result, three criteria are leveraged to filter the trajectories: 1) origin and destination within the scope of study; 2) the average speed of adjacent locations below 120km/h; 3) the time interval of adjacent locations less than 30 seconds. To generate road sequences, we further map the trajectory to the road network in OpenStreetMap using ST-Matching, which considers spatio-temporal constraints to map trajectory points with nearby roads. Due to the flexibility of travel, the transformed road sequence suffers from the inherent problem of inconsistent length. To this end, we compute the normalized frequency of the subsequence as a uniform representation of any two road sequences based on the idea of k-Gram. Compared to dynamic time warping, a method for calculating the similarity of unequal-length sequences, dividing the subsequence overwhelmingly reduces the runtime when dealing with a large number of trajectories (see Appendix A for the results). After many trials and error, we set the length of the subsequence to four by default, which can be modified through system interaction.

Route identification. We employ polar distance and hierarchical clustering to group similar sequences, considering them as a critical route. We conduct comparative experiments on different clustering methods. The results show that hierarchical clustering outperforms other methods (see Appendix B for the results) and is adopted in this work. This approach offers a distinct advantage in handling sparse vectors and eliminates the need for pre-defined cluster numbers. Because road network structure has a significant impact on travelling, we introduce the sequence identity metric to evaluate the clustering effect by measuring the overlap of routes.

For a single route, the objective is to maximize the matching score through Equation 1 when traversing a pair of road sequences, with $ms(s_\alpha^i, s_\beta^j) = E_a$ for an equal road, $ms(s_\alpha^i, s_\beta^j) = E_b$ for a non-equal road and $ms(s_\alpha^i, s_\beta^j) = E_c$ for inserting a gap to ignore a road, where E_a, E_b, E_c are constant terms. As a result, the sequence identity is the average of the percentage of roads that are identical.

$$\underset{ms}{\operatorname{argmax}} \sum_{\alpha=1}^{m_1} \sum_{\beta=1}^{m_2} ms(s_\alpha^i, s_\beta^j) \quad (1)$$

Pattern extraction. Given the road sequence of a route, we use the vertical maximum sequence pattern to identify typical travel patterns (i.e., road subsequence that frequently passed). Different from other methods, it only traverses the data once, significantly saving the cost of scanning data. After a trade-off between pattern coverage and execution time, we set the minimum support to half the number of trajectories by default.

4.2. Quantitative Analysis of Influencing Factors

Figure 3 illustrates the process of exploring influencing factors. The Ranking SVM model is trained to learn a function $g(X; W)$ (i.e., hyperplane) that separates the positive and negative instance. Utilizing Ranking SVM transforms the quantification of influencing factors into a ranking problem, addressing the issue of insufficient route utility information.

Derivation of preference. As described in Section 4.1, we extract multiple routes and use the influencing factors calculated in Section 3.1 as their feature. The input of Ranking SVM comprises a set of training instances X , where each instance consists of a pairwise difference vector of factor ($x_e = F_{R_i} - F_{R_j}$) and a label indicating the relative order ($y_e \in [1, -1]$). Initially, routes are ranked according to their popularity (i.e., the number of trajectories). Let the label be 1 if one route is more popular than the other, otherwise -1. Next, we feed X into the model and generate an optimized marginal function through Equation 2. Where, $\|*\|^2$ denotes the L_2 paradigm; η is the coefficient; $\forall \lambda_e > 0$ denotes the slack variable; (a, b) denotes the dot product of elements and J denotes the number of extracted routes. Once the model is trained, we can get the weight of different factors on route choice from an overall perspective, that is, $W = \{w_1, \dots, w_d\}$.

$$\begin{aligned} \min_{W, \lambda} \quad & \frac{1}{2} \|W\|^2 + \eta \sum_{e=1}^J \lambda_e \\ \text{s.t.} \quad & y_e(x_e, W) \geq 1 - \lambda_e \end{aligned} \quad (2)$$

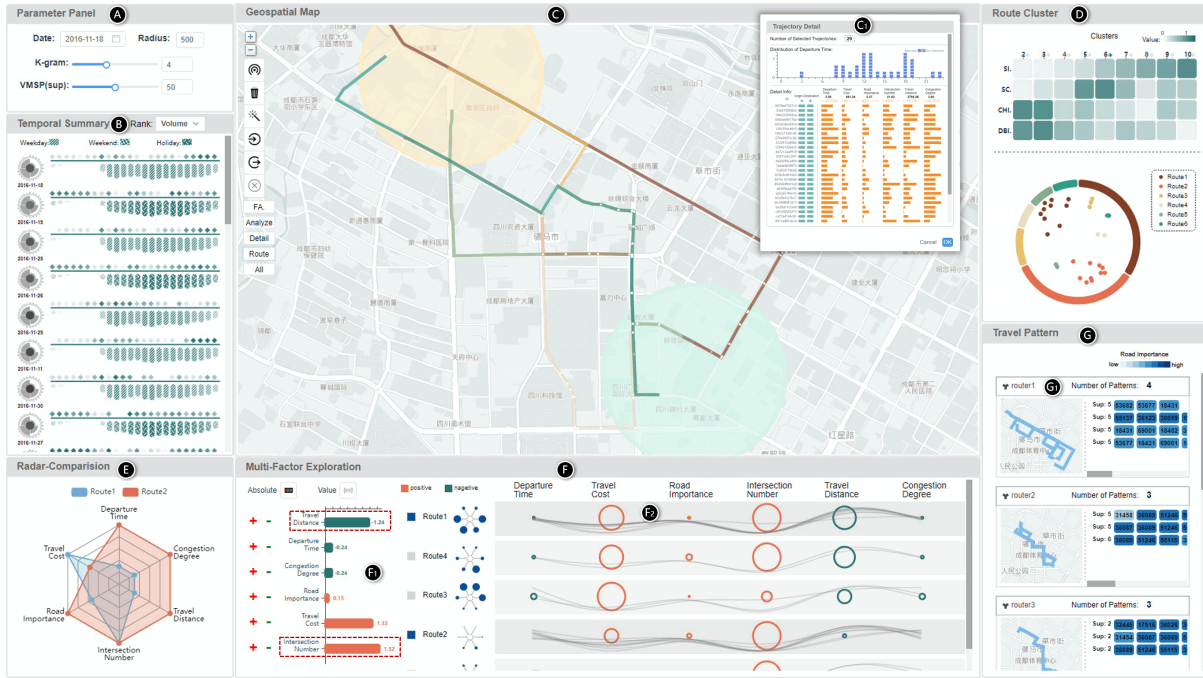


Figure 4: Visual interface of RouteVis. (A) Parameter panel is used to configure algorithm parameters. (B) Temporal summary view overviews the temporal information of the trajectory. (C) Geospatial map view shows the real status of people’s travels with trajectory detail (C1) and interactions. (D) Route cluster view presents the result of the identified routes and evaluation metrics. (E) Radar view and (F) multi-factor exploration view supports the analysis and comparison of the influencing factors on route choice. (G) Travel pattern view reveals the road subsequence where people travel frequently.

To help analysts understand and compare the role of factors in guiding route choice, we use a simple additive weighting method with excellent interpretability. According to Equation 3, the contribution is the product of weight and raw value with respect to the largest ones, where, $f_d^{R_u}$ denotes the value of the d_{th} factor in route R_u ; w_d is the corresponding weight and $\max_{d'}(f_{d'}^{R_u} \cdot w_{d'})$ is the maximum value of the product for normalization.

$$Con(w_d, f_d^{R_u}) = \frac{f_d^{R_u} \cdot w_d}{\max_{d'}(f_{d'}^{R_u} \cdot w_{d'})} \quad (3)$$

5. Visual Design

5.1. Design Requirements

Based on a questionnaire survey of 55 participants from different social groups (see Appendix C for the details), we summarize the following four design requirements.

R1: Generating a visual overview of the trajectory. The trajectory data involves spatio-temporal information, reflecting people’s daily travel patterns and route choice preferences. When exploring massive trajectories, it is necessary to provide a holistic overview, which not only follows the analysis principle of "overview first, zoom and filter, then details on demand [Shn96]", but also provides visual clues for users to drill down and select interesting trajectories for in-depth analysis.

R2: Clustering routes based on characteristics. There may be

many alternative routes to travel from a specific origin to a destination. It is not practical to consider all trajectories during analysis, as it is a time-consuming and labor-intensive process. For alleviating the cognitive burden of analysts, the system should have the capability to group similar trajectories, which can assist the summarization of travel patterns.

R3: Quantifying the effect of different factors on route choice. The route choice process can be influenced by a variety of factors and changes with different times and travel scenarios. Therefore, the visual representation should provide valuable insights into how people make travel decisions, which can inform transportation planning and policy-making.

R4: Enabling the flexible exploration of route choice for different travel contexts. A common scenario is that a person commuting to work during peak hour may prioritize the fastest route, however when traveling for leisure, the same person may prefer a scenic or non-congested route. Aiming to capture these differences, the system should provide sufficient interaction to analyze route choice in different contexts.

5.2. Temporal Summary View

The goal of the temporal summary view is to assist traffic analysts gain a holistic picture of daily travel, which serves as a foundation for data analysis (R1). Radial layouts excel at visualizing multidimensional information, whereas bar charts are well-suited for presenting discrete data. Therefore, we use these two charts to

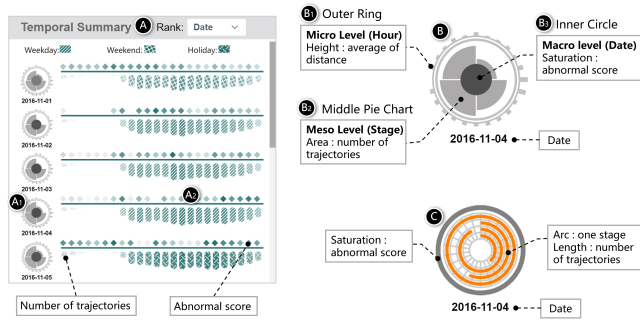


Figure 5: Temporal summary view. (A) Visual encoding of daily travel information. (B) Multi-level glyph. (C) Alternative design.

visualize the processed trajectory data. Each row represents a day and consists of a multi-level glyph (Figure 5(A1)) and a subview (Figure 5(A2)). All elements in the subview are arranged from left to right, which represents one hour between 0:00 and 24:00. The texture of each subview encodes the type of day, where the color saturation of each diamond and the height of each bar shows the abnormal score and the number of trajectories per hour. The abnormal score is measured by the number of trajectories per day or hour using the Isolation Forest algorithm. A multi-level glyph (Figure 5(B)) provides analysts with the status of daily travel, which facilitates the selection of the date. In the tailored glyph, three kinds of information are encoded, that is, the height of the outer ring shows the average distance per hour (B1); the area of the middle pie represents the number of trajectories at each stage (B2); the color saturation of the inner circle shows the abnormal score (B3).

Alternative. Figure 5(C) shows an early prototype of our design, where the length of the arc indicates the number of trajectories in each stage and the color saturation of the outermost circle encodes the abnormal score. However, the length of the arc is dependent on the radius, which can cause discrepancies and misinterpretations when expressing the same value. As a result, it is discarded.

5.3. Geospatial Map View

The spatial context is crucial for analysts because it enhances the perception of people's actual movements. To this end, a geospatial map view is used to display the geographic distribution of travel, while providing sufficient information to brush areas worth studying (R1, R2). Analysts can investigate regions of interest in two ways: 1) they can brush any area within a specific radius using the circle; 2) they can directly click on two functional areas in the map. As shown in Figure 6(A), if the analyst brushes or clicks the origin and destination, the system marks them with green and yellow, respectively. By default, all trajectories are dynamically displayed with blue lines accompanied by white dots. When the fifth button is clicked, each identified critical route is mapped with a qualitative color scheme. To alleviate visual clutter, we only mark the representative route (i.e., trajectory with the highest sequence identity score). In addition, we provide thirteen interactions (Figure 6(A1)) to support analysts in getting different aspects of information. Due to the high density of origins or destinations, we utilize a heatmap to visually encode them from red to blue, thereby overcoming the

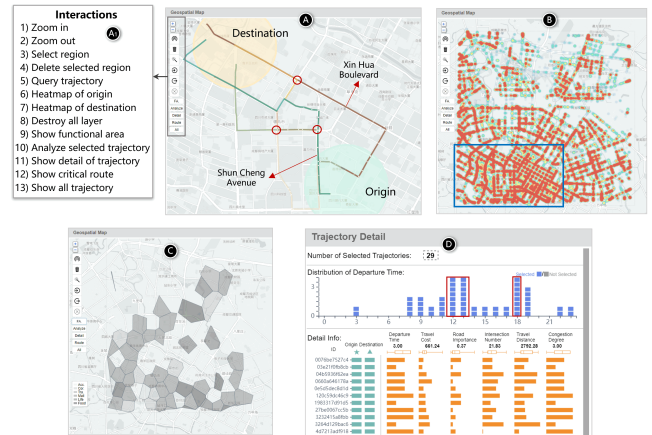


Figure 6: Geospatial map view. (A) Visual design and Interactions. (B) Heatmap of origin or destination. (C) Distribution of functional areas. (D) Detail of trajectories.

obstacle caused by point overlap. As shown in Figure 6(B), locations with more trajectories are represented in a deeper red.

Division of traffic functional area. The traffic functional area is a crucial component of the urban transportation network, and its layout and characteristics also affect route choices. Hence, we exploit Mini Batch K-Means to cluster the origin and destination, and determine the functionality according to the maximum number of points of interest. As a result, Cheng Du can be divided into 63 functional areas, represented in different shades of gray (Figure 6(C)), including 3 accommodation facilities, 9 corporations, 1 transportation center, 38 malls, 4 living services, and 8 food outlets.

Details of trajectory. The statistical results of departure time and influencing factors are temporal and discrete data, respectively. We design variants of bar charts for an intuitive visual representation (Figure 6(D)). In the first row, the dashed box at the top shows the total number of trajectories within the brushed area. In the second row, a stacked bar chart shows the distribution of the number of trajectories at each departure time, where the x-axis and y-axis indicate hour and the number of trajectories. If the analyst is interested in a certain departure time, further investigation can be done by clicking on the corresponding bar. In the last row, horizontal bars are designed to understand the characteristics of the influencing factors for each trajectory. We express their average values explicitly at the top, while drawing a horizontal whisker plot to display the statistical information. The length of the bar represents the normalized value of each factor.

5.4. Route Cluster View

It is a common method to analyze the clustering results by using the scatterplot with color. Based on this experience, we design a route cluster view (Figure 4(D), R2) which includes the evaluation (top) and the clustering results (bottom). The matrix shows the relationship between multiple variables in a grid format, which is ideal for the many-to-many relationship in the evaluation results. In the figure, each row represents a metric, and each column represents the total number of clusters formed. To facilitate a comparison, the

same metrics are normalized and encoded with the color saturation of a matrix cell. Additionally, analysts can click on the cross at the top of each column to interactively check the result of clustering. Aiming to reduce the cost of learning, the clustering result is integrated with a typical scatterplot and donut chart. A data point located inside represents a trajectory and is colored depending on the route to which it belongs. Note that the color scheme is consistent with the geospatial map view. The layout of the scatterplot is obtained through multidimensional scaling of polar distance. Finally, we summarize the number of trajectories contained in the route drawn as a donut chart. Each slice represents a route, and the size of the slice corresponds to its proportion of the whole. Analysts can click on each slice to hide alternative routes in geospatial map view and highlight the specific route of interest for tracking.

5.5. Multi-Factor Exploration View

After exploring the data using the proposed quantitative analytical framework, the results of analysis exhibit characteristics of multivariate (i.e., multiple influencing factors) and multi-level (i.e., the overall weight, contribution to individual route, and raw values) nature. Considering the learning curve for analysts, the multi-factor exploration view (Figure 4 F) provides a wealth of valuable information by ingeniously using simple visual elements such as line, circle, and color (R3, R4). Concretely, the length of the bar (Figure 7(A1)) describes the weights of the influencing factors, where green means negative and orange means positive. If the derived weights do not coincide with the analyst's prior knowledge, a plus/minus button is provided at the bottom of the parameters for modification. In the middle of this view (Figure 7(A2)), we visualize the multivariate information (i.e., the average of raw influencing factors) of the route in the form of an egocentric network arranged in a clockwise direction. Besides, the dashed circle in the center represents the route itself. The larger the radius of the circle, the larger its value. On the basis of ensuring information integrity and a concise user interface, the system provides mouse hovering, enabling analysts to quickly grasp the implicit meaning of each circle along with its corresponding actual numerical values (Figure 7(A5)). Furthermore, we depict the contribution of influencing factors through a tailored visualization that integrates the idea of matrix and parallel coordinates. Overall, each row represents a route, where the transparency of the gray background indicates the number of trajectories it contains, and each column indicates a factor (Figure 7(A3)). The radius of the circle indicates the contribution value, where the color is consistent with that in Figure 7(A1). Similar to the parallel coordinates, a line in each row illustrates the distribution of a trajectory over different factors, providing an intuitive overview.

Comparison of route. The analyst can select any two routes for comparison by clicking on the square in front of the route (Figure 4(E)). The radar chart is particularly useful when multiple variables need to be presented and compared simultaneously. Thus, we represent each factor as a separate axis in a clockwise manner and connect the data points of each route with a line, which enables the analyst to assess the relative differences at a glance.

Alternative. To facilitate visual perception, the initial scheme uses a superimposed matrix to display the data value and contribution (Figure 7(B)), where the row and column represent route and

factor, respectively. For any factor, the transparency of the background cell indicates the data value (i.e., the darker the color, the greater the value), while the size of the foreground cell indicates the contribution (i.e., the larger the cell, the greater the contribution). If the foreground cell completely covers the background cell, it difficult to identify the value accurately. For this reason, we make an improvement that uses the line to represent the data values.

5.6. Travel Pattern View

The travel pattern view, comprising a series of cards (Figure 4(G, G1)), presents the analysts with typical road subsequences mined from diverse trajectories by the vertical maximum sequence pattern algorithm (R2). To preserve the travel information, we still show the trajectory with a blue line on the mini-map. A row on the right represents a pattern consisting of support count and road subsequence. We visualize the subsequence with horizontally stacked rectangles, where the number maps the road ID and the transparency is determined by the road importance. Constrained by space, we provide a scrollbar to access content that exceeds the visible area.

5.7. Interactions

A collection of interactions (e.g. ranking, hovering, clicking, and brushing) is available for analysts to explore the dataset efficiently. A circular brush and functional area selection are provided on the geospatial map view and the trajectories that are not within the scope of the research are automatically filtered. In the multi-factor exploration view, RouteVis defaults to rank the factors by numerical magnitude, while providing rank by absolute value.

6. Evaluation and Discussion

6.1. Usage Scenario

Consider a traffic analyst whose routine is to report on the state of traffic at the end of the year. We describe how the proposed approach helps traffic analysts perform their duties by envisioning two typical examples.

Case1: Which route do people travel most often? (R1, R2, R3) Understanding the critical route helps transportation planners gain insights into vital roads within a traffic network, which can inform long-term planning efforts. The following examples demonstrate the effectiveness of the proposed method in extracting routes.

We initially observe daily travel information through the temporal summary view, as shown in Figure 4(B). A remarkable phenomenon is that the color of the leftmost bar is lighter and becomes darker from 7:00, reaching a peak at 13:00 or 14:00, which coincides with daily life. Ranked by the volume of the trajectory, we can see that the trajectories on November 18th are the most numerous, with a total of 187,912. Then, we select that day in the parameter panel (Figure 4(A)) for further exploration, and switch to the heatmap of origin to observe the spatial distribution. Relative to other areas, the QingYang District (blue rectangle in Figure 6(B)) has higher traffic volumes.

We take Evergrande Central Plaza and QingYang Government as the origin and destination (Figure 6(A)) to explore how people

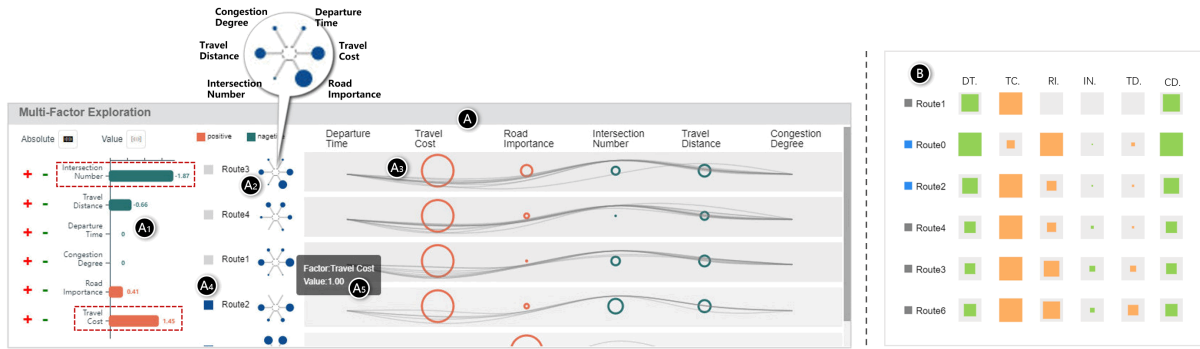


Figure 7: Visualization of influencing factors. (A) Multi-factor exploration view. (B) Alternative design.

travel. After setting the exploration radius to 500 meters in the control panel, we brush the aforementioned areas with the lasso tool. It can be seen from the Figure 6(D) that there are 29 trajectories in total, with peaks at 12:00, 13:00, and 18:00 (red rectangles). Observing the characteristics of the influencing factors, it can be seen that most of them set off in the third stage (i.e., 12:00 18:00), while the distance and duration are short, at 2792 meters and 661 seconds. It is notable that the average congestion is 3 (i.e., mildly congested with a speed between 20km/h and 30km/h).

In the route cluster view, the system automatically identifies 10 critical routes because of the highest sequence identity score. Since the silhouette coefficient is the largest in 6, we believe that identifying 6 routes is more suitable. Intuitively, from the scatterplot of the route cluster view (Figure 4(D)) we can see that Route 1 (colored in brown) and Route 2 (colored in orange) are the more popular routes, each with 10 trajectories. Additionally, the geospatial map view visually displays that all the routes pass through Shun Cheng Avenue, except for Route 1 which passes along Xin Hua Boulevard (red arrow in Figure 6(b)). Obviously, the routes traveling in this area form two junctions and one divergence marked with a red circle in Figure 6(A), which provide clues as to how to optimize traffic flow. It is interesting that Route 4 takes a detour to the destination at Shu Du Boulevard. Further, in the travel pattern view (Figure 4(G)), we note that the color of the mined frequent patterns is dark blue, which indicates that people usually choose to travel on trunk, while Road 36,089 and Road 51,246 have more traffic.

Case2: Why do people prefer a particular route? The route on which people choose to travel can depend on various factors such as location and time, which poses a challenge for traffic analysts. We verify how the system reveals the preferences of people's route choice for different times and travel purposes.

Comparison of different travel time (R4). For the aforementioned areas, we subsequently observe the weights of the influencing factors derived from the Ranking SVM model in the multi-factor exploration view. Overall, the length of travel distance, travel cost, and intersection number is larger, meaning they have the greatest influence on route choice. Specifically, travel distance has a negative effect (-1.24, Figure 4(F1)), suggesting that people prioritize routes with a shorter distance, which coincides with our daily observations. The weight of the intersection number is positive (1.52, Figure 4(F1)), presumably due to the fact that these vehicles

travel in busy urban centers. This phenomenon is also supported by the travel cost, whose weight is positive, implying that people travel for a long duration, despite the short distance. Further research is needed to explore whether this is caused by congestion or unreasonable signal control.

Next, we inspect the contribution of each factor to route in detail. We can see that Route 1 is the best route, followed by Route 4 and 3 (Figure 4(F2)). Consistently, the radius of the circle on the left is larger with respect to travel cost, intersection number, and travel distance, indicating a greater contribution. Upon further inspection, it turns out that the high ranking of Route 1 is mainly attributed to the intersection number, although its average travel distance is much larger compared to other routes. Since there are also 10 trajectories in Route 2, we click on the square of Route 1 and Route 2 to perform a one-to-one comparison in the radar chart (Figure 4(E)). Both have an equal value in terms of intersection number, while the difference in travel distance, congestion degree, and departure time ultimately results in Route 1 being ranked ahead.

To understand the differences in influencing factors over time, we continue to analyze the six identified route on November 19th. As shown in the multi-factor exploration view, the obvious difference is that the intersection number has the greatest negative impact (-1.87), which is the opposite of November 18th (Figure 7(A1)). We believe that people prefer a route with fewer intersection number, but pay little consideration to congestion degree and departure time. We also note that travel cost is still a more important positive factor (Figure 7(B1)), in other words, the duration of the commute between Evergrande Central Plaza and QingYang Government is longer at any time. Traffic managers can implement adaptive signal control systems, as well as upgrade and expand road infrastructure to reduce travel duration and improve the travel experience.

Differences in the purpose of travel (R4). An individual's purpose of traveling from one place to another can influence their priorities and preferences when selecting a route. We compare and analyze three typical scenarios (namely travelling from accommodation to a corporation, a mall, and a transportation center) to demonstrate that the proposed system supports the exploration of route choices in various contexts.

In the geospatial map view, we switch to the functional area by clicking the FA button. Taking the Xi Hua Men Community as the origin, we study the travel to the Province City Government Affairs

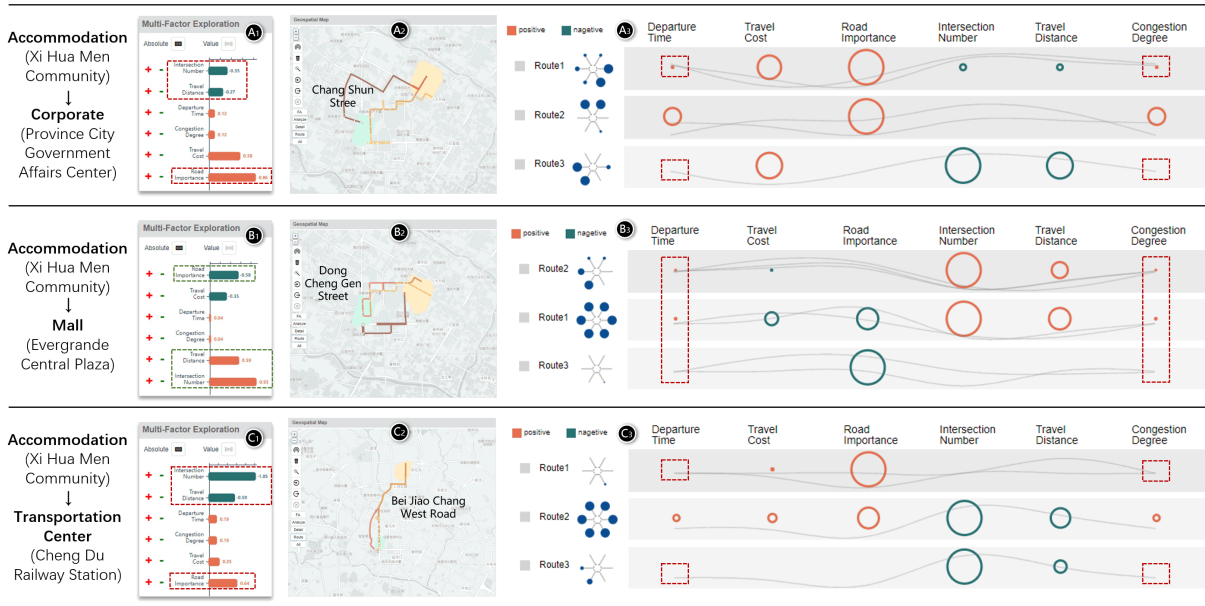


Figure 8: Exploration and comparative analysis of influencing factors in different contexts.

Center (corporation), Evergrande Central Plaza (mall) and Cheng Du Railway Station (transportation center), respectively. In general, these travels are strongly influenced by road importance, intersection number, travel distance, and travel cost, as the length of the corresponding bar in Figure 8(A1, B1, C1) is longer. Furthermore, the influence of departure time and congestion degree is relatively weak, so the circles on them is also fewer and smaller (red dashed-line rectangle in Figure 8(A3, B3, C3)).

Note that there are several differences as well. Firstly, people are more concerned about the convenience and efficiency of travel from their accommodation to a corporation and a transportation center. As shown in Figure 8(A1, C1), when people travel from their accommodation to a corporation or a transportation center, intersection number (-0.35, -0.27) and travel distance (-1.05, -0.59) are negative influences.

Second, road importance has the strongest positive effect on route choice, i.e., 0.86 and 0.64, regardless of whether it is from accommodation to a corporation or a transportation center. With the help of the geospatial map view, we learn that people are driving on a primary or secondary road such as Chang Shun Street (Figure 8(A2)), Dong Cheng Gen Street (Figure 8(B2)), Bei Jiao Chang West Road (Figure 8(C2)), etc.

Finally, the preference of route choice from the accommodation to the mall is completely the opposite to that of travelling to corporation and transportation center (Figure 8(B1)). Road importance has a negative influence (-0.58), while intersection number and travel distance have a positive influence (0.93 and 0.59). A reasonable explanation is that people are more inclined to consider comfort when traveling without time constraints, even though these routes may involve a longer distance or more intersections.

6.2. Expert Interview

We invite three domain experts to evaluate the system regarding workflow, visual design, and effectiveness. E1 and E2 are employees of a government transport department and an enterprise, respectively. They have been closely collaborating for a long time to explore people's travel patterns and analyze the current status of urban transportation operations. E3 is a PhD student working in the field of transport with a foundation in visualization. None of the three experts have known about the RouteVis system before.

Procedure. We conducted a two-hour online meeting with three experts. First, we spent 10 minutes introducing the RouteVis system, including the data, addressed problems, workflow, and analytical requirements. Subsequently, we explained the visualizations and interactions of the RouteVis system using a demo video of usage scenarios in Section 6.1. Finally, we asked experts to freely explore the system and collect their feedback during usage, providing constructive guidance for further improvement.

System. The comprehensive and clear analytical process of RouteVis had been praised by three experts. Experts only needed to select areas of interest in the geospatial map view, and the system would automatically explore. As stated by E1, "The RouteVis system makes the entire analysis process easy and efficient. Compared to the tedious operations of spreadsheet queries and statistics, it significantly enhances our work efficiency". E1 favored the multi-factor exploration view, as he believed that this approach was practical in the analysis of influencing factors, providing compelling evidence for decision-making. However, E1 also pointed out that "it would be better if the findings from the exploration could be documented". E2 emphasized the organic synergy among various views, "Combining multi-factor exploration view and radar view, I can more flexibly compare the differences in influences on different routes". In addition, E2 suggested, "Since the system has al-

Table 2: Running time (milliseconds) for each step.

Trajectory Number	k-Gram	Similarity	Clustering	VSMP	Ranking SVM
17	9.97	446.78	2905.24	36.33	9.97
23	1.00	515.62	1996.36	14.67	4.99
53	5.98	2139.28	6008.63	16.33	4.99
77	5.99	5716.70	20979.87	17.00	4.99
115	27.92	8393.54	25917.78	30.00	10.97
156	48.90	17271.76	55758.42	34.33	3.99
207	71.79	41385.26	199398.28	56.00	3.99
350	58.87	86424.72	117913.45	37.33	5.97
445	153.59	157374.93	260784.31	51.00	24.93
793	242.35	385683.82	277310.82	53.67	3.99

ready provided the travel pattern view, should it be considered as a factor?". Both E1 and E2 recommended providing downloadable guidelines to describe how to use the system.

Visual Designs. Expert E1, E2, and E3 all reported that the system provided richer information through diverse visual channels. Regarding the temporal summary view, they thought it's good to show information at different levels of granularity, i.e., days, stages, and hours. E3 commented, "After watching the demo video, I could clearly understand the meaning of each graphical element and color scheme in the system". A trade-off between visual representation and cognitive load was required. "RouteVis requires a certain learning curve, especially for novice users," added E3. Besides, E1 questioned the lines showing the raw data in the multi-factor exploration view. "These lines allow me to observe the overall trend, but I struggle to obtain specific details without the y-axis".

7. Discussion

System Performance. The running time of the RouteVis system includes the computation of the quantitative analytical framework and the rendering of the visualization. We recorded the required running time for each step by conducting experiments on 10 randomly selected travel data, as shown in Table 2. As the number of trajectories increases, the response of k-Gram, VMSP, and Ranking SVM is rapid, taking only tens of milliseconds. The bottleneck of RouteVis lies in similarity calculation and clustering. Theoretically, the computational complexity of similarity is $O(n^2)$, which depends on the number of trajectories n . To ensure that the user experience is not affected by prolonged waiting times, we employ interactive filtering within the system, so as to exclude trajectories that are not expected to be analyzed. Currently, RouteVis has hundreds of trajectories after filtering, so route identification can be done in a few minutes.

Generalization. Trajectory data can be regarded as spatio-temporal data because it contains information in both spatial and temporal dimensions, so the methods and visual designs proposed in this paper can be applied to explore other data. The core task of analyzing influencing factors can be abstracted as: identify the objective, define the factors, and quantify the weight. For example, a similar exploration scheme can be used to discover the causes of air quality pollution. The analysis process involves identifying contaminated areas (i.e., clustering monitoring data) and deter-

mining whether they are caused by industrial or vehicle emissions. The spatio-temporal characteristics of air pollution are summarized through the geospatial map view and temporal summary view in RouteVis. Potential influencing factors are then calculated and quantified, using statistical or machine learning methods. The multivariate presentation solution (e.g. egocentric network and matrix) in the multi-factor exploration view still is applicable.

Limitations. The scalability of route identification is somewhat constrained, primarily due to two factors: the computation of similarity and the traversal to find the optimal route. Given the spatiotemporal characteristics of the transportation domain, offline route recognition demonstrate significant effectiveness in reducing user waiting times. However, it suffers from a lack of flexibility in meeting user analytics requirements. Moreover, the scalability of the system is also a concern for us. We only allow users to explore up to 10 routes. On the one hand, the spatial representation of the road in the map is limited, and displaying too many routes can result in overlap, making it challenging to discern the actual driving process. Although we provide interactions allowing analysts to track a single route, mitigating visual clutter to some extent. On the other hand, in the route cluster view, using the qualitative color schemes to visually represent identified routes faces scalability issues. As the number of routes increases, the distinction between colors becomes smaller and smaller. The other limitation is that the functionality of the system needs to be further improved. RouteVis should save a snapshot of the analysis process to facilitate backtracking to the previous state and compare the results and differences at different stages.

8. Conclusion and Future Work

In this work, we study the quantification of influencing factors that affect route choice. Based on summarized design requirements, we propose RouteVis, a visual analytics system to help traffic analysts understand the preferences of people's route choice behind trajectory data. Through trajectory transformation, we capture the similarity between them and utilize hierarchical clustering to identify routes while evaluating them with four metrics. We then propose a quantitative approach to model route choice behaviors and derive their respective influence weights. Usage scenarios and expert interviews demonstrate that RouteVis can effectively discover the relationship between route choice and the influencing factors and enhance traffic analysts' perception of the current state of travel.

However, there is still room for improvement in this work: 1) We will strive to incorporate additional factors for a more comprehensive and detailed understanding of people's decision-making processes. The potential effects of external factors such as weather, major events, or infrastructure changes on route choice cannot be ignored. 2) An in-depth study on how multiple influences work together on route choice is an optional direction. We look forward to revealing the interrelationships among these influencing factors and their comprehensive effects on route choice behavior.

Acknowledgements

This work was supported by National Natural Science Foundation of China under Grant 42171450 and 62377008, the Key Re-

search and Development Project of Science and Technology Development Plan of Jilin Province under Grant 20210201074GX and 20240302071GX.

References

- [11520] 115-2020 G.: *Evaluation Method for Road Traffic Congestion Levels*. The Ministry of Public Security of the People's Republic of China, Beijing, China, 2020. 3
- [AAC*17] ANDRIENKO G., ANDRIENKO N., CHEN W., MACIEJEWSKI R., ZHAO Y.: Visual analytics of mobility and transportation: State of the art and further research directions. *IEEE Transactions on Intelligent Transportation Systems* 18, 8 (2017), 2232–2249. 2
- [ABWL98] ADAMOWICZ W., BOXALL P., WILLIAMS M., LOUVIERE J.: Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *American Journal of Agricultural Economics* 80, 1 (1998), 64–75. 1
- [ADKZ*17] AL-DOHUKI S., KAMW F., ZHAO Y., MA C., WU Y., YANG J., YE X., WANG F., LI X., CHEN W.: Semantictraj: A new approach to interacting with massive taxi trajectories. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 11–20. 2
- [AHS15] ALIVAND M., HOCHMAIR H., SRINIVASAN S.: Analyzing how travelers choose scenic routes using route choice models. *Computers, Environment and Urban Systems* 50, 2 (2015), 41–52. 2
- [BAB99] BEN-AKIVA M., BIERLAIRE M.: Discrete choice methods and their applications to short term travel decisions. *International Series in Operations Research and Management Science* 23 (1999), 576–582. 2
- [BZA17] BAPAT S. V., ZALA D. L. B., A.A.AMIN: Route choice analysis: A case study of nh 8 and ne1. *International Journal of Recent Trends in Engineering and Research* 3, 5 (2017), 216–227. 1
- [BZQ*21] BAI J., ZHANG H., QU D., LV C., SHAO W.: Fgvis: visual analytics of human mobility patterns and urban areas based on f-glove. *Journal of Visualization* 24, 6 (2021), 1319–1335. 2, 3
- [CGW15] CHEN W., GUO F., WANG F.-Y.: A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems* 16, 6 (2015), 2970–2984. 2
- [CK22] CHO S.-H., KHO S.-Y.: Heterogeneous route choice model incorporating group segmentation based on travel experience. *KSCE Journal of Civil Engineering* 26, 3 (2022), 1376–1387. 2
- [CXL*06] CAO Y., XU J., LIU T.-Y., LI H., HUANG Y., HON H.-W.: Adapting ranking svm to document retrieval. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (2006), pp. 186–193. 2
- [CZKZ22] CHEN Z.-G., ZHAN Z.-H., KWONG S., ZHANG J.: Evolutionary computation for intelligent transportation in smart cities: A survey. *IEEE Computational Intelligence Magazine* 17, 2 (2022), 83–102. 2
- [DS77] DAGANZO C. F., SHEFFI Y.: On stochastic models of traffic assignment. *Transportation Science* 11, 3 (1977), 253–274. 2
- [DS17] DALUMPINES R., SCOTT D. M.: Determinants of route choice behavior: A comparison of shop versus work trips using the potential path area - gateway (ppag) algorithm and path-size logit. *Journal of Transport Geography* 59 (2017), 59–68. 1
- [HMK*20] HERBERTH R., MENZ L., KRPER S., LUO C., GAUTERIN F., GERLICHER A., WANG Q.: Identifying atypical travel patterns for improved medium-term mobility prediction. *IEEE Transactions on Intelligent Transportation Systems* 21, 12 (2020), 5010–5021. 2
- [HZC*20] HUANG Z., ZHAO Y., CHEN W., GAO S., YU K., XU W., TANG M., ZHU M., XU M.: A natural-language-based visual query approach of uncertain human trajectories. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 1256–1266. 2
- [LLY*17] LU M., LAI C., YE T., LIANG J., YUAN X.: Visual analysis of multiple route choices based on general gps trajectories. *IEEE Transactions on Big Data* 3, 2 (2017), 234–247. 2
- [LSD18] LU W., SCOTT D. M., DALUMPINES R.: Understanding bike share cyclist route choice using gps data: Comparing dominant routes and shortest paths. *Journal of Transport Geography* 71 (2018), 172–181. 1
- [LWL*17] LIU D., WANG D., LI Y., BAO J., ZHENG Y., QU H., WU Y.: Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 1–10. 2, 3
- [LWY15] LU M., WANG Z., YUAN X.: Trajrank: Exploring travel behaviour on a route by trajectory ranking. In *2015 IEEE Pacific Visualization Symposium* (2015), pp. 311–318. 2
- [MSL*19] MARKOVI N., SEKULJA P., LAAN Z. V., ANDRIENKO G., ANDRIENKO N.: Applications of trajectory data from the perspective of a road transportation agency: Literature review and maryland case study. *IEEE Transactions on Intelligent Transportation Systems* 20, 5 (2019), 1858–1869. 2
- [Pra09] PRATO C. G.: Route choice modeling: past, present, and future research directions. *Journal of Choice Modelling* 2, 1 (2009), 65–100. 2
- [PS11] PAPINSKI D., SCOTT D. M.: A gis-based toolkit for route choice analysis. *Journal of Transport Geography* 19, 3 (2011), 434–442. 2
- [PYSJ21] PI M., YEON H., SON H., JANG Y.: Visual cause analytics for traffic congestion. *IEEE Transactions on Visualization and Computer Graphics* 27, 3 (2021), 2186–2201. 2
- [QHGF19] QI G., HUANG A., GUAN W., FAN L.: Analysis and prediction of regional mobility patterns of bus travellers using smart card data and points of interest data. *IEEE Transactions on Intelligent Transportation Systems* 20, 4 (2019), 1197–1214. 2
- [Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages* (1996), pp. 336–343. 5
- [SJK*23] SHIN D., JO J., KIM B., SONG H., CHO S.-H., SEO J.: Rcmvis: A visual analytics system for route choice modeling. *IEEE Transactions on Visualization and Computer Graphics* 29, 3 (2023), 1799–1817. 2
- [SLB21] SCOTT D. M., LU W., BROWN M. J.: Route choice of bike share users: Leveraging gps data to derive choice sets. *Journal of Transport Geography* 90 (2021), 102903. 2
- [SZZ*14] SUN D., ZHANG C., ZHANG L., CHEN F., PENG Z.-R.: Urban travel behavior analyses and route prediction based on floating car data. *Transportation Letters: the International Journal of Transportation Research* 6, 3 (2014), 118–125. 2
- [TWH*20] TANG J., WANG Y., HAO W., LIU F., HUANG H., WANG Y.: A mixed path size logit-based taxi customer-search model considering spatio-temporal factors in route choice. *IEEE Transactions on Intelligent Transportation Systems* 21, 4 (2020), 1347–1358. 1
- [War52] WARDROP J. G.: Road paper. some theoretical aspects of road traffic research. In *Proc. Institute of Civil Engineers* (1952), vol. 1, pp. 325–362. 2
- [XZX11] XU H., ZHOU J., XU W.: A decision-making rule for modeling travelers's route choice behavior based on cumulative prospect theory. *Transportation Research Part C Emerging Technologies* 19, 2 (2011), 218–228. 2
- [YB20] YAO R., BEKHOR S.: Data-driven choice set generation and estimation of route choice models. *Transportation Research Part C Emerging Technologies* 121 (2020), 102832. 2
- [ZCX*19] ZHU M., CHEN W., XIA J., MA Y., ZHANG Y., LUO Y., HUANG Z., LIU L.: Location2vec: A situation-aware representation for visual exploration of urban locations. *IEEE Transactions on Intelligent Transportation Systems* 20, 10 (2019), 3981–3990. 2
- [ZFA*16] ZENG W., FU C.-W., ARISONA S. M., ERATH A., QU H.: Visualizing waypoints-constrained origin-destination patterns for massive transportation data. *Computer Graphics Forum* 35, 8 (2016), 95–107. 2

- [ZFAQ13] ZENG W., FU C.-W., ARISONA S. M., QU H.: Visualizing interchange patterns in massive movement data. *Computer Graphics Forum* 32, 3 (2013), 271–280. [2](#)
- [ZMT*19] ZHOU Z., MENG L., TANG C., ZHAO Y., GUO Z.: Visual abstraction of large scale geospatial origin-destination movement data. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 43–53. [2](#)
- [ZWC*16] ZHANG Y., WU W., CHEN Y., QU H., NI L. M.: Visual analytics in urban computing: An overview. *IEEE Transactions on Big Data* 2, 3 (2016), 276–296. [2](#)
- [ZWC*18] ZHANG H., WANG R., CHEN B., HOU Y., QU D.: Dynamic identification of urban functional areas and visual analysis of time-varying patterns based on trajectory data and pois. *Journal of Computer-Aided Design and Computer Graphics* 30, 19 (2018), 1728–1740. [2](#)