

Article

Visual Analysis Method for Traffic Trajectory with Dynamic Topic Movement Patterns Based on the Improved Markov Decision Process

Huarong Chen ¹ , Yadong Wu ^{2,*}, Huaquan Tang ³, Jing Lei ³, Guijuan Wang ¹, Weixin Zhao ⁴, Jing Liao ¹, Fupan Wang ¹ and Zhong Wang ¹

¹ School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China; chenhuarong@swust.edu.cn (H.C.); gui Juanwang@swust.edu.cn (G.W.); liaojing@swust.edu.cn (J.L.); wangfp@swust.edu.cn (F.W.); zhongw@mails.swust.edu.cn (Z.W.)

² School of Computer Science and Engineering, Sichuan University of Science & Engineering, Yibin 644007, China

³ Technical Center, Mianyang Xincheng Engine Co., Ltd., Mianyang 621000, China; huaquan.tang@xce.com.cn (H.T.); jing.lei@xce.com.cn (J.L.)

⁴ College of Computer Science, Sichuan University, Chengdu 610065, China; zhaoweixin@stu.scu.edu.cn

* Correspondence: wuyadong@suse.edu.cn

Abstract: The visual analysis of trajectory topics is helpful for mining potential trajectory patterns, but the traditional visual analysis method ignores the evolution of the temporal coherence of the topic. In this paper, a novel visual analysis method for dynamic topic analysis of traffic trajectory is proposed, which is used to explore and analyze the traffic trajectory topic and evolution. Firstly, the spatial information is integrated into trajectory words, calculating the dynamic trajectory topic model based on dynamic analysis modeling and, consequently, correlating the evolution of the trajectory topic between adjacent time slices. Secondly, in the trajectory topic, a representative trajectory sequence is generated to overcome the problem of the trajectory topic model not considering the word order, based on the improved Markov Decision Process. Subsequently, a set of meaningful visual codes is designed to analyze the trajectory topic and its evolution through the parallel window visual model from a spatial-temporal perspective. Finally, a case evaluation shows that the proposed method is effective in analyzing potential trajectory movement patterns.

Keywords: trajectory topic model; Markov Decision Process; visual analysis



Citation: Chen, H.; Wu, Y.; Tang, H.; Lei, J.; Wang, G.; Zhao, W.; Liao, J.; Wang, F.; Wang, Z. Visual Analysis Method for Traffic Trajectory with Dynamic Topic Movement Patterns Based on the Improved Markov Decision Process. *Electronics* **2024**, *13*, 467. <https://doi.org/10.3390/electronics13030467>

Academic Editor: Dah-Jye Lee

Received: 29 December 2023

Revised: 17 January 2024

Accepted: 20 January 2024

Published: 23 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Visual exploration of traffic trajectory data is constrained by the limited capacity for visual information and the lack of interactive guidance. Traffic trajectories record the historical travel of vehicles, with visual information and potential features at spatial-temporal levels. Exploring traffic trajectories through visualization can intuitively obtain visual information about the trajectories and mine the potential movement patterns. However, drawing the traffic trajectory directly on a map will be limited by the information capacity and will cause problems such as visual occlusion. The most intuitive method to help the visual exploration of traffic trajectory is to reduce the visual density by computing the trajectory features. Liu et al. [1] effectively reduce the visual density of Origin to Destination (OD) flow by dividing and exploring urban functional areas. Liu et al. [2] employ a tensor decomposition algorithm to segment multi-dimensional spatiotemporal data and reduce the number of visualization elements on the same screen. Zhou et al. [3] propose a visual exploration system to reduce the visual clutter and strengthen the relevance of OD flows. Deng et al. [4] provide an efficient visual presentation design of temporal events by abstracting the workflow of temporal and spatial interaction association analysis. Andrienko et al. [5] designed a method to abstract the characteristics of spatio-temporal

OD mobility to reduce the intersection and occlusion of the trajectory flow graph. Unfortunately, these methods calculate trajectory patterns at the data feature level and do not provide enough semantic guidance for the analyst. In addition, analysts usually need to manually select and iterate different motion data slices to search for hidden patterns in trajectory data, making it difficult to deeply analyze potential trajectory movement patterns from a spatial-temporal perspective. We address these issues and propose a set of meaningful visual encodings. This visualization model establishes a parallel window model. The model presents the trajectory topic characteristics from the spatial and temporal perspectives, which is helpful in reducing visual interference and providing guidance for visual interactive exploration.

The dynamic topic information in the trajectory topic model reflects the evolution of the topic. A topic model, which provides topic guidance for trajectory analysis by applying the topic model to trajectory data analysis, is a kind of statistical model. Most of the existing methods that apply topic models to trajectory modeling are based on Latent Dirichlet Allocation (LDA) [6], Non-negative Matrix Factorization (NMF) [7], and other unsupervised topic algorithms, focusing on optimizing topic extraction strategies. Specifically, Chu et al. [8] introduce the LDA to infer hidden patterns of moving taxi populations. Liu et al. [9] transform the traffic data into a document library to capture latent traffic semantic patterns based on the NMF model. Liu et al. [10] employ the bigram topic model to analyze textualized trajectories. Tao et al. [11] introduce a hierarchical topic model H-NMF to extract multi-granularity traffic topics to capture mobility patterns. Traffic trajectory is dynamic, and the corresponding trajectory topic constantly evolves. However, the trajectory topic analysis based on the above modeling methods only considers the trajectory spatial information. It does not consider the temporal dynamic evolution of the trajectory topic, resulting in the loss of trajectory topic evolution information.

Trajectory features are described by trajectory semantics [12]. Trajectory order is an essential part of the semantics of trajectory topics. The topic model describes the topic content by counting word frequency but does not consider the ordinal relationship between words. Traffic trajectories are sequentially connected trajectory segments with a sequential relationship. Trajectory sequence information is necessary for trajectory topic modeling. However, applying the topic model to the topic modeling of traffic trajectory is challenging and needs to overcome the influence of the topic model not containing word sequence information on the interpretability of trajectory topics. In summary, there needs to be more trajectory order semantics and topic evolution information in existing work on trajectory topics. This implies the need for visual encodings that provide sufficient semantic guidance to analyze trajectory topic evolution.

To solve these problems, a visual analysis method is proposed to analyze the dynamic topics of trajectories. This method considers the evolution information of trajectory topics and overcomes the problem of missing word sequence information in probabilistic models. Firstly, the dynamic topic model (DTM) [13] for a trajectory is established to analyze the trajectory topic evolution fitted by parameter distribution. Secondly, by calculating the thematic entropy, we can quantify the informational content of the trajectory topics, which indicates meaningful thematic time segments, thereby reducing the time required for the manual repetitive selection of data slices. Subsequently, we employ Markov chains to represent trajectory words and address the issue of topic models not retaining word order information. Based on the Markov chains, we utilize an improved Markov Decision Process (MDP) to generate trajectory sequences that represent the topics, which aids in analyzing the topic content. Moreover, we introduce a meaningful visual encoding scheme, presenting trajectory topic visualizations separately through parallel views from both abstract maps and thematic information perspectives. Lastly, a case study based on a real dataset was conducted to demonstrate the efficacy of the methodology proposed in this paper in exploring dynamic trajectory topics. In summary, the primary contributions of this work are as follows:

- A trajectory dynamic topic model is proposed. This model considers the topic evolution and is used to process traffic trajectory data. This model helps to obtain trajectory dynamic topics and capture their evolution. The topic entropy is calculated to represent the volume of information in trajectory topics and is utilized to analyze the evolution of the topic.
- The improved Markov Decision Process is proposed to generate representative trajectory sequences of a topic. This method aims to overcome the problem of probabilistic models which often lose word sequence information.
- The visual models, featured with the parallel and regular arrangement, are designed to help users explore the trajectory dynamic topics and related indicators.
- Experimental studies based on the publicly available dataset demonstrate the effectiveness of this method in trajectory dynamic topic exploration.

2. Related Works

2.1. Trajectory Topic Modeling

The basic idea of topic modeling is to vectorize a given corpus by term frequency or inverse document frequency. Topic modeling decomposes a document-term matrix into document-topic and topic-word subsets, subsequently optimizing these subsets using probabilistic or decomposition techniques. The clustering of implicit semantic structures in documents obtained through topic modeling can be applied to the discovery of trajectory semantic structures and trajectory topics. Trajectory topic modeling aims to describe motion patterns by calculating sub-trajectory frequency information, thereby further extracting the content of trajectory topics. Commonly used topic models include LDA [6], NMF [7], and others.

As common topic computing methods, LDA and NMF are widely used in topic modeling. LDA computes the posterior distribution of the data based on Bayesian algorithms, whereas NMF produces deterministic outcomes from matrix decomposition, not probabilistic values. Current efforts in topic modeling largely revolve around modifications or advancements of the LDA and NMF methodologies. For instance, Liao et al. [14] established a trajectory topic model based on LDA to unearth latent driving patterns and analyze a driver's travel intentions. Huang et al. [15] employed LDA to explore hidden ship mobility patterns from the trajectory document corpus. In contrast, Liu et al. [2] applied tensor decomposition to multi-dimensional spatiotemporal data to extract regional patterns. Addressing the challenge of selecting an optimal number of topics, Liu et al. [9] proposed an interactive trajectory topic modeling method based on a semi-supervised non-negative matrix factorization (SS-NMF) for trajectory data. Similarly, Tao et al. [11] presented a hierarchical topic analysis approach to balance the analyses of varying granularities of movement patterns. Notably, NMF's performance degrades with more significant numbers of topics, whereas LDA exhibits more robust modeling results.

Beyond using LDA and NMF for trajectory topic modeling, Liu et al. [10] adopted the Biterm Topic Model (BTM) [16] to model trajectory movement patterns, leveraging the Apriori algorithm for sub-trajectory extraction and using sub-trajectory sequences to indicate directionality. Originating for short-text analysis, BTM differs from LDA in that its word generation is corpus-wide. Without a document hierarchy, BTM primarily retrieves word probabilities related to topics, making it challenging to obtain document-specific topic probabilities. However, different trajectory topics are implied in trajectory datasets. Analyzing trajectory data deeply and obtaining specific topic probabilities using BTM is difficult.

These topic modeling methods, including LDA, NMF, and BTM, calculate the trajectory topic within the time interval and do not calculate the trajectory topic evolution between time slices at the model level. Dynamic topic modeling adds a time factor, and the topic content changes with time. Corresponding Dynamic topic versions exist for NMF and LDA, namely, Dynamic NMF [7] and DTM [13]. Dynamic NMF [17] uses the state space to represent the NMF time dependence. The dynamic NMF first divides the corpus into time

slices and then runs the NMF algorithm on the time window to obtain the topic of each time slice. After finding topics in each time slice, the dynamic NMF classifies semantically similar topics as dynamic topics. DTM is based on LDA but uses variational methods to connect topic variations in different time slices. Compared with Dynamic NMF, DTM has more continuity between adjacent time slices.

A significant limitation of these methods is their focus on trajectory topics within specific time intervals without accounting for temporal evolution at the model level. Dynamic topic modeling introduces time as a variable, allowing topic content to evolve. NMF and LDA have their dynamic counterparts, specifically Dynamic NMF and DTM. Dynamic NMF represents NMF's temporal dependencies using state spaces, segments the corpus into temporal slices, and then runs the NMF algorithm over these time windows. After identifying topics within each time slice, semantically similar topics are categorized as dynamic topics. On the other hand, DTM, grounded in LDA, employs variational methods to link topic variations across distinct temporal slices, offering enhanced continuity compared to Dynamic NMF. Yao et al. [18], leveraging non-trajectory data from social media, developed a DTM for spatiotemporal sequence analysis. Specifically, they achieved dynamic topic analysis by covering adjacent temporal slices with radii from social media locational points. This approach bypasses trajectory point connections and computes trajectory topics at the individual point level.

In general, there is an issue of weak continuity in trajectory topics across slices when establishing a trajectory topic model. To address this problem, this paper introduces DTM into the analysis of dynamic trajectory topics for the first time, continuously tracking the evolution of trajectory topics through DTM. Additionally, by improving the Monte Carlo method to represent transition probabilities between trajectory segments, it constructs trajectory segment sequence information, providing more details for describing the evolution of trajectory topics.

2.2. Visual Analysis of Trajectory Topic

Visual analysis is often used to explore the inherent features and patterns encoded in trajectory data, because of the connection between humans and computers provided by interactive visual interfaces [19,20]. Topic visualization primarily revolves around the interrelationships among topic sets, encompassing the temporal evolution of topics and the static relations between them. The Theme River [21], an early visualization approach, is well-suited for analyzing topic evolution. While the fluid visual characteristics of the Theme River facilitate capturing change trends over time, its river-like curve metaphor poses challenges for quantitative analysis of topic attributes. In addition, Sankey diagrams are another standard visualization tool. They use rectangles to quantify topic evolution but can become crowded and complex when there is a multitude of topics. To overcome the above problems, the parallel matrix heatmap is employed for quantitative analysis of topic evolution, enabling the identification of trends in topic evolution. The visual feature of its neat arrangement allows for a clear and intuitive interpretation of visual information.

Methods for trajectory topic visual analysis merge geographical and visualization views, offering multi-level spatial information for trajectory topic pattern analysis. He et al. [22] incorporated LDA into trajectory topics to depict spatio-temporal topic variations, harnessing visual filters to mine spatio-temporal text data. Similarly, Chu et al. [8] employed LDA to dissect vast amounts of taxi trajectory data within cities, unveiling concealed knowledge through visualization. Similarly, Gao et al. [23] leveraged LDA to extract topics related to road cyclists' behavior and movement semantics, adopting clustering and interactive visualization techniques to identify further clusters of similar cyclists and road segments within the city. By visualizing and retrieving taxi trajectory data, AL-Douhuki et al. [24] provide a semantic management approach for taxi trajectory data. Ali et al. [25] proposed a topic modeling approach grounded in LDA based on ontology, extracting topics and features by retrieving traffic content from social networks. Gao et al. [23] used LDA to extract the topics of cyclist behavior semantics and moving semantics on roads. Then,

they implemented multiple interactive visual interfaces to facilitate the interpretation for analysts. Yan et al. [26] utilized a BTM to unearth latent topics within trajectories and subsequently analyze these topics via visualization.

Concerning trajectory points, Zhou et al. [27] and Liu et al. [1] implemented sampling strategies to reduce visual clutter, facilitating the discovery of trajectory patterns. Recognizing the large-scale nature of trajectory data, Zhou et al. [3] designed visual representations of Origin-Destination (OD) flows, iteratively probing OD movement patterns. Andrienko et al. [5] aggregated data to identify spatial similarities in trajectories and abstracted spatio-temporal features to delve into trajectory flows. Wang et al. [28] introduced a multi-scale geographical area data visualization method predicated on spatial attribute associations, enhancing visual perception of both global and detailed features. While these methodologies provide visual optimization at the trajectory point and line levels, investing considerable effort in visually exploring trajectory features, they fall short in contrasting spatial trajectory topic information and topic features between temporal slices.

In summary, to effectively explore the evolution of trajectory topics from different perspectives of trajectory points, trajectory lines, and trajectory topics using visual components, this paper proposes a multi-entity cascade visual analysis method under parallel views. This method aims to further reveal the potential patterns of dynamic trajectory topics by presenting the topic features through parallel abstract graphs and parallel heat matrices containing time series information. The parallel view overcomes the challenges of quantitatively analyzing topic river graphs and dealing with the complexity of crowded Sankey diagrams by organizing visual elements in a parallel layout and cascading the analysis of topic semantics. This study studies the construction and characteristics of dynamic spatio-temporal trajectory graphs to comprehend the evolution of dynamic trajectory topics.

3. Methods

3.1. Overview

One of the challenges in trajectory topic modeling and analysis is effectively integrating both temporal and spatial information into the topic computation. To address these intricacies, our study employed three specific strategies.

First, concerning temporal information, we utilized trajectory segments partitioned based on specific time intervals to model dynamic trajectory topics. Given that the starting and ending points of trajectory segments are fixed, there is no ambiguity in trajectory words, meaning each trajectory word has a singular interpretation. Furthermore, the MDP generates trajectory sequences within the collection of trajectory segments representing a trajectory topic, ensuring that spatially discrete trajectory segments are interconnected. By leveraging these three strategies, we can efficiently compute trajectory topics imbued with both temporal and spatial dimensions. In the following sections of this chapter, we will delve deeper into these methodologies, providing a comprehensive exposition of their intricacies and applications.

3.2. Trajectory Dynamic Topic Modeling

Topic modeling employs statistical methods to discern abstract topics within documents. When applied to trajectory information, it facilitates the extraction of abstract trajectory topics. The DTM, in contrast to conventional topic models such as LDA and NMF, incorporates temporal sequence information, making it particularly well-suited to the temporal variability inherent in trajectories. A study closely aligned with our approach is [9], where the authors employed NMF to topic-model trajectories across different time periods. However, such an approach lacks temporal coherence in trajectory topics. Compared to NMF, DTM offers superior coherence and yields more robust and flexible fitting results.

DTM treats the posterior distribution of the model parameters in the current time window as the conditional distribution for the parameters in the subsequent time window.

Building upon the foundation of LDA by integrating a temporal dimension, DTM becomes especially adept at analyzing topic evolution. However, being a probabilistic model, DTM overlooks word order during its computational process. Such an oversight can be problematic, given that word order mirrors the sequential nature of trajectory segments. Our study addresses this limitation through an enhanced MDP, which we will delve into comprehensively in Section 3.3.

The trajectory DTM describes the trajectory motion pattern with time. In the DTM of traffic trajectory, we take the trajectory segment as a word and integrate the trajectory spatial information into the word. The trajectory data is divided according to time slices, with the latter moment evolving from the previous moment. Since the Dirichlet distribution is not suitable for use in time series models, as shown in Equation (1), a Gaussian noise evolved state space model is used to connect the natural parameter β about the trajectory topic to sample the word distribution $\Phi(Z(d_n))$ about the trajectory topic $Z(d_n)$. In Equation (2), the natural parameter α about the trajectory topic ratio is also connected using the Gaussian noise evolution state-space model to sample the trajectory topic distribution $\Theta(d)$ of the generated trajectory document d . In the DTM model, the prior probabilities α for document topics and the distribution β for topics are dynamically iterated.

$$\beta_{t+1,k} | \beta_{t,k} \sim N(\beta_{t,k}, \sigma^2 I) \quad (1)$$

$$\alpha_{t+1,k} | \alpha_{t,k} \sim N(\alpha_{t,k}, \delta^2 I) \quad (2)$$

Equations (1) and (2) link the natural parameters β_t and α_t for each topic, where β_t and α_t are the Gaussian noise natural parameters for time slice t and topic k . σ and δ are the variance parameters. After linking both the trajectory topic and its proportion distribution, we sequentially bind the collection of trajectory topic models. The corpus generation process for the trajectory DTM is as follows: Within time slice t , a possible word (trajectory segment) distribution β for trajectory topic k is first sampled. Then, a potential topic distribution $\alpha(t)$ for each document (a collection of trajectory segments) is sampled. For each document within time slice t , the topic distribution η for document d is obtained. For every word in the document during time slice t , the possible topic $z(t, d, n)$ for the n th word in document d is sampled, and the corresponding word $w(t, d, n)$ is generated. The number of topics is determined using an algorithm that identifies the optimal topic count based on perplexity. This generation process's graphical model is illustrated in Figure 1. Both α and β at time t are generated based on α and β at time $t - 1$. Using time dynamics, the k th topic at slice t smoothly evolves from the k th topic at slice $t - 1$. When the horizontal arrow is removed, the graphical model is reduced to a set of separate topic models.

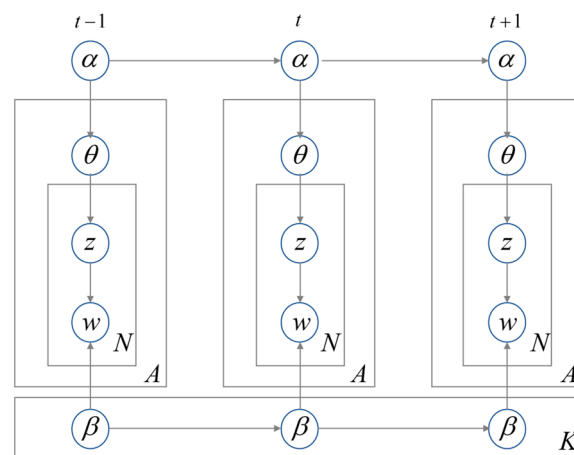


Figure 1. The generation process of the trajectory dynamic topic model includes three time slices [15].

The goal of the trajectory DTM is to calculate the posterior distribution. The objective of the variational method is to optimize the free parameters of the distribution over latent variables, making the distribution closely approximate the true posterior in terms of Kullback–Liebler (KL) divergence and subsequently substituting it for the genuine posterior. Within DTM, a variational approach based on the variational Kalman filter approximation is utilized to compute the posterior distribution [13]. Entropy measures the average amount of information expressed in every allocation of a random variable, depicting the system's level of disorder. We posit that when the system exhibits a low level of disorder, the trajectory topic is relatively stable. Conversely, with high disorder levels, the trajectory topics are relatively diverse. We compute the information entropy concerning topic probabilities within each time slice to provide information quantity across varying time slices, assisting in the subsequent selection of representative time slices for further analysis. The topic entropy at time t can be expressed as in Equation (3), where p represents the topic probability distribution at moment t .

$$H(t) = -\sum_{i=1}^K [p(t) \log p(t)] \quad (3)$$

3.3. Trajectory Dynamic Topic Generating

3.3.1. Sub-Trajectory Modeling

The trajectory dynamic topic consists of sub-trajectory sets that adhere to Markovian properties. Extracting representative trajectory sequences from these sub-trajectory sets embodies the trajectory topic. Trajectory segments illustrate transitions between states. Following a particular policy, the MDP can generate trajectory sequences within the dynamic topic sub-trajectory sets that adhere to the Markovian properties, representing the trajectory topic. A Markov chain depicts the random process of transitioning from one state to another and is a mathematical construct composed of a set of random variables. Sub-trajectory sets describe inter-regional transitions, fundamentally representing state transitions. We can characterize the transition relationships within these sub-trajectory sets using Markov chains. The trajectory topic is composed of the sub-trajectory set $E = \{e_1, e_2, \dots, e_n\}$. The set of areas for trajectory transitions form the state space $S = \{s_1, s_2, \dots, s_n\}$. A sub-trajectory x_i , transitioning from region s_i to another region x_j , exhibits frequency characteristics. According to Markovian properties, the probability of moving from the current state to the next is only dependent on the current state, and not on prior locations, adhering to the following Equation (4):

$$\Pr(X_{n+1} = x | X_1 = x, X_2 = x_2, \dots, X_n = x_n) = \Pr(X_{n+1} = x | X_n = x) \quad (4)$$

where $\{X_1, X_2, \dots, X_n\}$ represents the random state sequence comprised of state space S . $P_r(X_{n+1} = x | X_n = x)$ represents the probability of the current state X_n transitioning to the next state X_{n+1} . Given that the random process of the trajectory topic is a mathematical construct regarding the set of random variables of the sub-trajectory combinations, we can derive representative trajectories within the trajectory topic by constructing a Markov chain related to the sub-trajectory topic.

The Markov chain uses a probabilistic automaton to display transition probabilities between states, resulting in the transition matrix P . This matrix is a series of directed graphs, with edges representing the probability of transitioning from S_n to S_{n+1} . Each state in the state space is included once as a row and a column, with the matrix indicating the probability of moving from a row state to a column state and the probabilities of each row summing up to 1. The steady-state distribution of the Markov chain, regarding the sub-trajectory set, aids in generating consistent trajectory sequences. Within the sub-trajectory set, the time-invariant Markov chain $\{X_n, n \geq 0\}$ of the transition matrix P has its initial state space S_i , conforming to the probability distribution $\varnothing = \{\varnothing_0, \varnothing_1, \dots, \varnothing_n\}$, satisfying the matrix Equation (5):

$$\varnothing = \varnothing P \quad (5)$$

where $\emptyset$ represents the steady-state distribution of the Markov chain, meaning a time-invariant Markov chain is wholly described by its initial state space S_i and its transition matrix P . We simulate the trajectory topic's initial state probability distribution $\emptyset_i$ and transition matrix against matrix Equation (5). As illustrated in Figure 2, the initial state of the sub-trajectories within the topic satisfies the stationary distribution, which can make the Markov process of the trajectory topic a stationary random process. As shown in Figure 2, the vertical axis is the probability of sub-trajectory state, and the horizontal axis is the number of iterations. Different colored line segments represent different trajectory segments. After approximately 60 simulation steps, a steady state is achieved under the action of the transition matrix P , with the distribution beginning to converge to a stable distribution at a certain point. Hence, the initial state of the sub-trajectory collection concerning the trajectory topic complies with the steady-state distribution. This implies that no matter the initial state of the trajectory, the equilibrium distribution will always be the same, achieving a stable initialization. The MDP concerning the sub-trajectory becomes a stationary stochastic process, meaning its finite-dimensional distributions are translation-invariant. This can further generate stable Markov sequences that represent the trajectory topic.

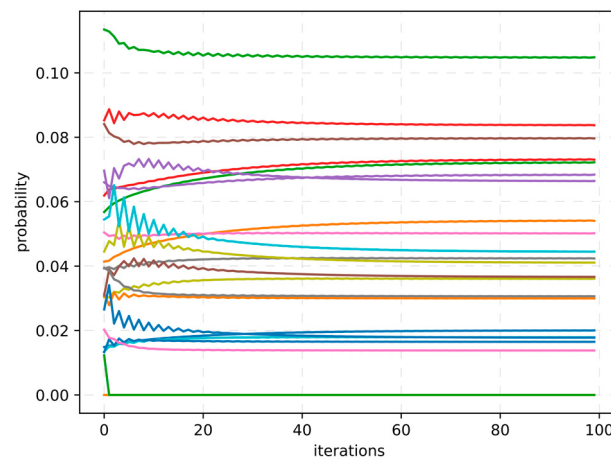


Figure 2. The trend of transition probability of sub-trajectories within a trajectory topic as the number of iterations increases. Lines represent trajectory topics, distinguished by different colors.

3.3.2. Trajectory Sequence Generating

The MDP elucidates the decision-making aspect of state transitions within Markov sequences. Within the sub-trajectory sets of the trajectory topic, transitions between sub-trajectories adhere to Markovian properties. Representative trajectory sequences within the trajectory topic can be generated using the MDP. This sequentially decided representative Markov sequence can serve as a succinct representation of the trajectory topic.

Formally, the MDP concerning the sub-trajectory set can be represented by the tuple $\langle S, A, R, P, \pi \rangle$. The state $S = \{s_1, s_2, \dots, s_n\}$ signifies the set of regions for trajectory transitions. The action space $A = \{a_1, a_2, \dots, a_n\}$ depicts transitions between states, essentially forming the starting point S_i and endpoint S_j of a sub-trajectory. The reward R reflects the reward achievable when transitioning from one state to another within the action space. As illustrated by Equation (6):

$$R = (1 + k)P - kL_p \quad (6)$$

the reward value R is composed of two components, specifically the transition probability P and the penalty term Φ . The penalty term Φ is composed of hyperparameters k and L_p .

Our aspiration is for the target sub-trajectory segment to maintain a relatively consistent speed, fluctuating within a certain speed range. As presented in Equation (7):

$$L_p = |V - V_{bm}| \times \left(\sum_{i=0}^N v_i - v_{bm_i} \right)^{-1} \quad (7)$$

the objective of establishing the penalty term L_p is to curtail trajectory segments with substantial speed deviations. Consequently, we obtain the time slices of the trajectory dynamic topic and the preset number of trajectory segments N to compute the baseline speed V_{bm} . Subsequently, trajectory segments with significant deviations are penalized. The transition probability $P : S \times A \times S \rightarrow [0, 1]$ expresses the probability of an action transitioning from one state to another. The essence of the policy π is to maximize the reward for each action. Specifically, π commences from a stable steady-state initial condition. During each trajectory transition, it randomly selects actions based on the transition probability, conducts finite dynamic lookups, and eventually attains the Markov sequence representing the trajectory topic.

3.4. Data Set and Hyperparameter Settings

We chose taxi trajectory data from Chengdu as our experimental dataset for our study. This dataset comprises approximately 14,000 taxi-formatted trajectories commonly used for trajectory research. The dataset includes vehicle ID, latitude, longitude, passengers, date, and time attributes from 3 August 2014, to 30 August 2014 [29]. Before running the trajectory dynamic topic modeling, we calculated, on an hourly basis, the hyperparameters suitable for the DTM on this data. We employed perplexity and coherence to guide the acquisition of the optimal topic number. A lower perplexity indicates better topic performance. Coherence computation measures the consistency of words constituting a topic. A higher coherence score means topic words reinforce each other, enhancing the interpretability of the topic. Figure 3 describes the consistency score results of the experimental dataset. After evaluating the performance of all measurement metrics, we settled on 17 as the number of topics to be computed.

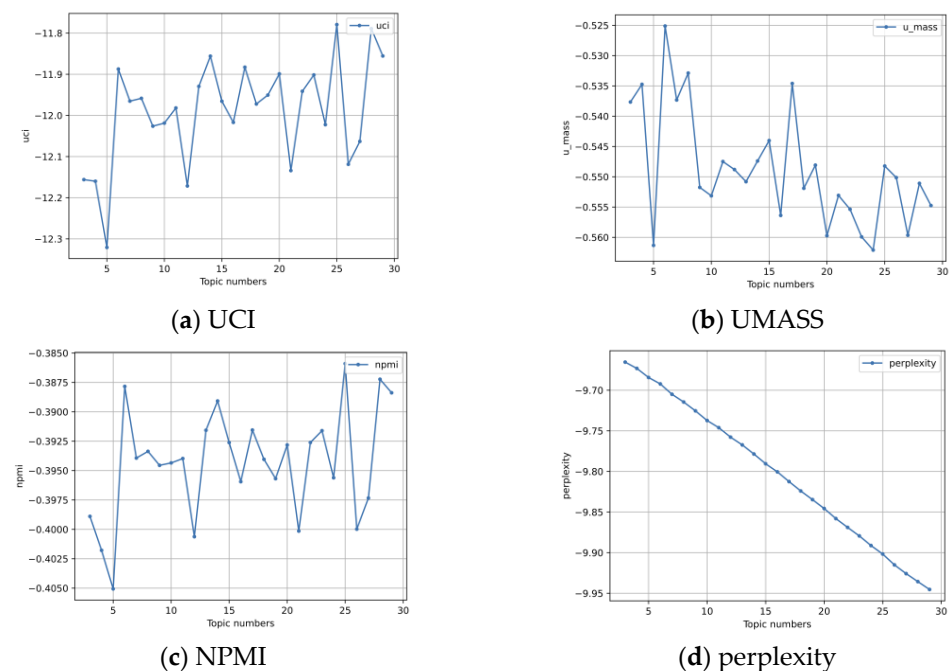


Figure 3. The coherence scores and perplexity are calculated which describe topic performance. Topic coherence measures the coherence of words within a topic; higher coherence is preferable. Indicators describing topic coherence include UCI [30], UMass [31], and NPMI [32]. Lower perplexity indicates a better topic model.

4. Visual Design

From a spatiotemporal perspective, we aim to dissect the composition and evolution of trajectory topics. The visual analysis of dynamic trajectory topics presents challenges. First, the dynamism of trajectory topics requires an intuitive representation through visualization, showcasing sequential changes across time and space. Secondly, given that dynamic trajectory topics are expressed as a series of trajectory word sets, a need arises to envisage the three-dimensional sequential evolution information linking time intervals, trajectory topics, and topic words.

We introduce the Parallel Viewport Visualization method for dynamic trajectory topic analysis to address these challenges. This technique provides a clear depiction of the dynamic shifts in trajectory topics through its parallel temporal visualization module and abstract map correlation module. This section delves deeper into the visualization technique for analyzing dynamic trajectory topic models.

4.1. Trajectory Dynamic Topic Time Evolution Analysis

The evolution of trajectory topics necessitates a compelling portrayal of their progression over time, offering a visual glance of the topic components. The information regarding the evolution of dynamic trajectory topics is embedded across multiple temporal slices, spanning various topics, each bearing its temporal evolution of topic words. This calls for an efficient visualization encompassing the three-dimensional data related to time, topic, and topic words. We calculated the topic distribution for each time slice and the evolution of specific topics between different time slices. We further assessed the information content of time slices and topics using entropy calculations. As illustrated in Figure 4, trajectory dynamic topic calculation consists of three parts. First, the original trajectory data is processed into a time-slice document composed of trajectory words, and then the trajectory dynamic topic is obtained by DTM and improved by MDP. Finally, the spatial and temporal evolution information of the trajectory dynamic topic is analyzed through the map view, parallel abstract map, and parallel matrix heatmap.

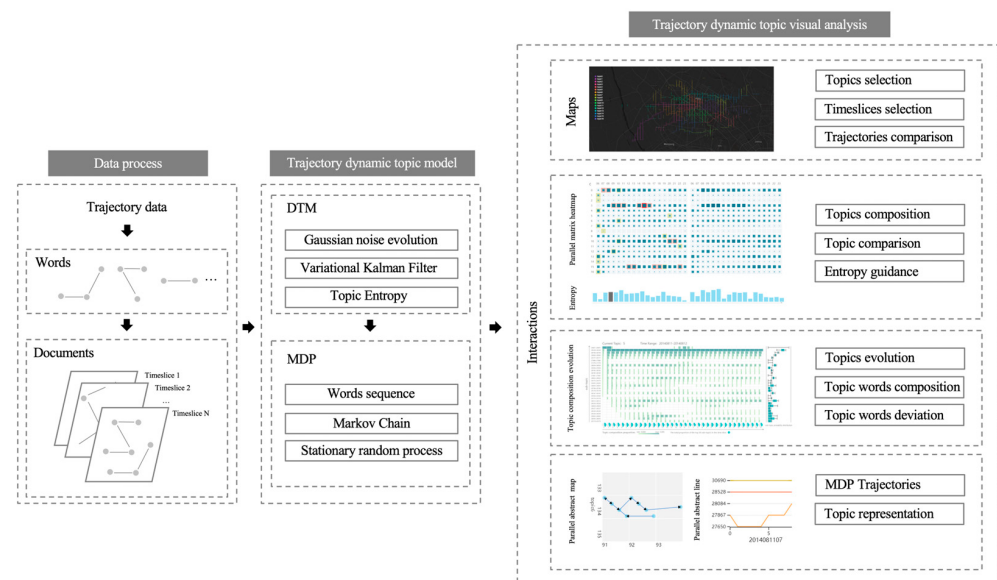


Figure 4. Calculation pipeline of trajectory dynamic topic. It represents the data analysis process from processing trajectory data and establishing a dynamic trajectory model to visual analysis of the dynamic trajectory model. Starting with data processing, the dataset is partitioned based on time slices, forming trajectory documents composed of trajectory segments. Subsequently, the DTM is utilized to establish a dynamic trajectory topic model. MDP is employed to generate representative trajectory sequences. Finally, combining maps and parallel views, spatial and temporal visual analysis is conducted to explore the evolution of trajectory topics.

While topic river plots are a conventional tool to visualize topic transitions, their curve estimations can introduce visual uncertainties. Thus, it becomes essential to sidestep visual biases stemming from visual encoding. To overcome these problems, we adopted the parallel matrix heatmap on cascade analysis. Comprising two parallel heatmaps and associated components, it facilitates efficient visual searches of dense three-dimensional information within a visually organized layout.

Figure 5C's topic evolution heatmap and Figure 5E's topic word heatmap form the parallel matrix heatmap, offering a concise and intuitive representation of time-topic and time-topic word information. The hierarchical construction of the time-topic and time-topic word cascade analysis allows for interaction across different hierarchical levels. In Figure 5C's heatmap, the x -axis denotes time, and the y -axis represents topics, with rectangle sizes encoding the composition of trajectory topics for corresponding time slices. To optimize exploration, we outlined rectangles with the highest probabilities in each row and column to swiftly pinpoint representative sub-topics. To reduce user interaction time when selecting the target trajectory's time slice, we computed the entropy of topics for each time slice, aiding users in choosing meaningful topic time slices, and displayed entropy values using bar charts on the right side of the heatmap. Similarly, the entropy of each topic was computed for target topic selection and was represented by using bar charts above the heatmap.

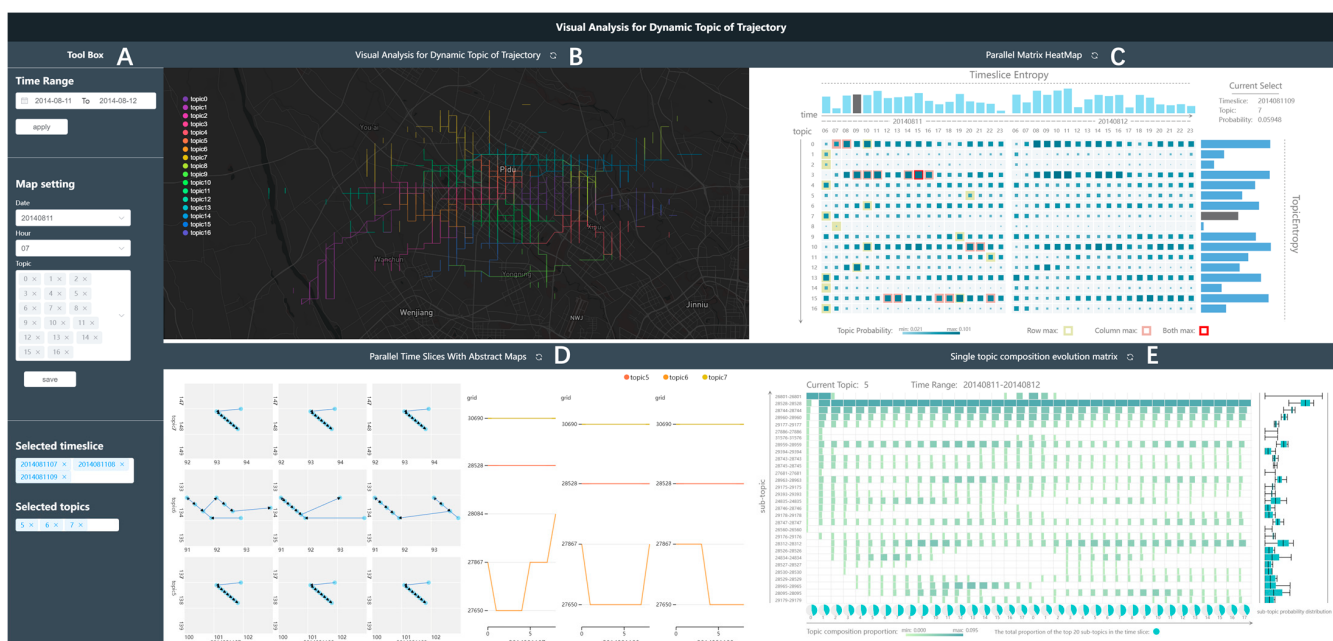


Figure 5. The developed trajectory dynamic topic visualization system includes visual components: toolbox, trajectory topic map, topic evolution heatmap, parallel abstract map, and topic word heatmap. These visual components correspond to (A–E) respectively. The system shows the dynamic topic of trajectory from 11–12 August 2014, in one-hour time slices.

Within the dynamic trajectory topic, topic words constituting the topic would evolve over time. As depicted in Figure 5E's topic word heatmap, the evolution of the topic components is described by the changing proportions of topic words. Specifically, the heatmap elements were shifted from the center-expanding rectangles to left-aligned, right-expanding rectangles to facilitate clearer vertical comparisons among topic words within the same category. Typically, the x -axis of the topic word heatmap represents time, and the y -axis represents the topic words constituting the topic. For this study, we chose to display the top 35 topic words based on probability rankings, with their respective proportions among all topic words represented by a pie chart below. One can select topics of interest

from the topic evolution heatmap to update the topic word heatmap, thus achieving cascade analysis concerning time, topics, and topic words.

4.2. Spatial Evolution Analysis of Trajectory Dynamic Topic

Exploring trajectory topics through maps provides a spatial perspective on the features inherent to such trajectories. A trajectory topic is not just a solitary line but also a complex mosaic comprising numerous sub-trajectory segments. It is much like deciphering the narrative of a tale that unfolds across different periods in time, spotting the changes in topics between time slices, and comparing the alterations within these sub-trajectory segments. In our quest to extract the essence of these trajectory topics, we have crafted parallel abstract maps. Just imagine running the MDP on sub-trajectory segments within time slices. This allows us to gauge transition probabilities, eventually leading us to a trajectory sequence that truly captures the spirit of the overarching trajectory topic. This spatial data is about geographical coordinates and the sequence in which they are arranged. Abstract maps convey the trajectory topics by depicting the spatial relative positions of trajectory points. Utilizing parallel windows in abstract maps allows for presenting the evolutionary information of trajectory topics from both temporal and spatial dimensions. This offers an intuitive comparison of representative trajectory sequences across different periods and topics.

As illustrated in Figure 5D's parallel abstract map, the horizontal axis represents time, while the vertical axis signifies topics. Users can select their areas of interest in the topic evolution heatmap to update the parallel abstract map, facilitating an analysis of trajectory topic evolution. This multi-instance comparison of spatial information effectively highlights the differences among trajectory sequences. Furthermore, to counteract the loss of spatial sequence information due to the overlapping trajectories on the map, we utilize a topic trajectory segment line chart to represent the trajectory topic sequence details. This chart conveys the order and composition of trajectory segments related to specific topics, offering insights into the relationship between trajectory sequences and their corresponding topics. Additionally, users can select topics of interest within the parallel abstract map to refresh the actual map, delving deeper into the detailed information of the trajectory topics.

5. Case Study

In the process of modeling transportation trajectory data, we adopted a trajectory DTM that considers topic evolution. This facilitates the extraction of dynamic trajectory topics and their subsequent evolutions. The target audience for this dynamic trajectory topic analysis includes those keen on understanding regional vehicle traffic patterns, such as urban planners and optimization specialists.

To demonstrate the efficacy of our methodology in the domain of dynamic trajectory topic analysis, we conducted a case study based on taxi data from Chengdu. This data spans 28 days, yielding 5,184,637 trajectory sequences after data cleaning. The granularity of analysis is correlated with the number of topics; the greater the number of topics, the finer the details within each topic. For this case, we defined 17 topics, with a minimum of 6 iterations for LDA, and set α to 0.01 to compute the dynamic trajectory topics.

5.1. Trajectory Topic Evolution Analysis

There are noticeable differences in trajectory topics within the same time slice. A single time frame encompasses multiple trajectory topics, each exhibiting varying intensities. Figure 5B illustrates the trajectory topic map at 7 a.m. on 11 August 2014, where most topics exhibit localized distributions. Having discerned the spatial distribution of trajectory topics, it is imperative to delve deeper into topic evolution over time.

5.1.1. Spatial Evolution of Trajectory Topics

There are significant differences between the trajectory segments that form trajectory topics at different time slices in the map. Before comparing the spatial evolution charac-

teristics of the trajectory topics, we need to determine the target time slice. In Figure 5C, the topic evolution heatmap offers a vertical perspective, revealing different topic probabilities within identical time slices. The entropy for time slices 07 and 23 on 11 August 2014, is relatively low, signifying minimal differentiation among the topic probabilities and, consequently, a smaller information content. In contrast, the entropy for time slices 11 and 15 on the same date is markedly higher, with apparent disparities in topic probabilities, translating to more extensive information content.

The composition and location of trajectory segments within different trajectory themes vary across time slices. Based on the discussion above about the entropy of time slices, we select the 07 time slice to examine the spatial evolution of trajectory topics. As depicted in Figure 6, a comparative analysis of topics 3 and 12 over consecutive time frames demonstrates an uptrend in trajectory segment count with increasing topic probability.

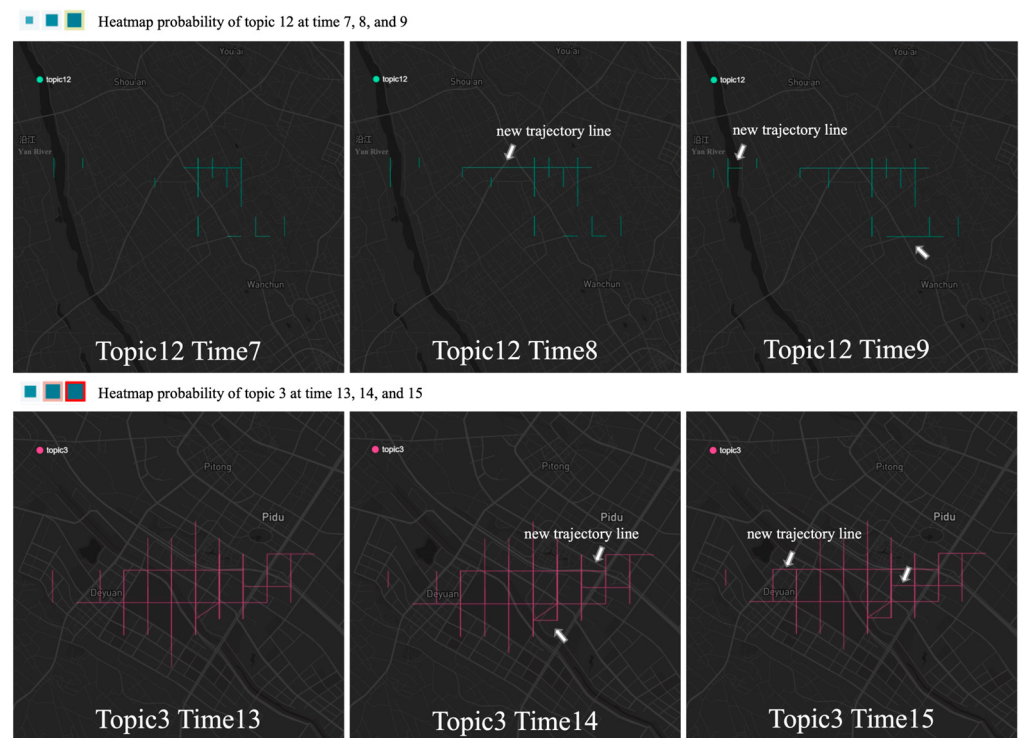


Figure 6. Trajectory topics are spatial and evolve over time. The trajectory segments that constitute a trajectory topic are subject to change. The evolution of trajectories for topic 3 and topic 12 on the map across three time slices is displayed. The white arrow indicates the new trajectory line. Trajectory topic 12 adds the middle trajectory line at time 8 and the left and lower right trajectory lines at time 9. Trajectory topic 3 adds the next volume and the right trajectory line at time 14 and adds the left, middle, and right trajectory lines at time 15.

5.1.2. Trajectories Topic Time Evolution

The information entropy of trajectory topics shows distinct differences across different time slices. One can examine the evolution of specific topics by observing Figure 5C's heatmap from a horizontal perspective. The bar chart on the right of Figure 5C shows the information entropy of topic evolution. Topic entropy is used to indicate the amount of trajectory topic information. Topic 8 can be seen to have the smallest information entropy because topic 8 has the smallest probability and little evolutionary difference. Topics 3 and 15 have high entropy due to the significant change in topic probability during the evolution of these topics. Further, we can analyze the trajectory topic evolution at the topic word level. Figure 7 shows the composition evolution of trajectory words between trajectory topic 8 and trajectory topic 10 from “20140811” to “20140812”. The topic entropy of trajectory topic 8 is smaller than that of trajectory topic 10. That is, the change degree

of topic words of trajectory topic 8 is smaller than that of trajectory topic 10. This also indicates that the number of trajectory words required to form trajectory topic 8 is smaller than those required for trajectory topic 10.

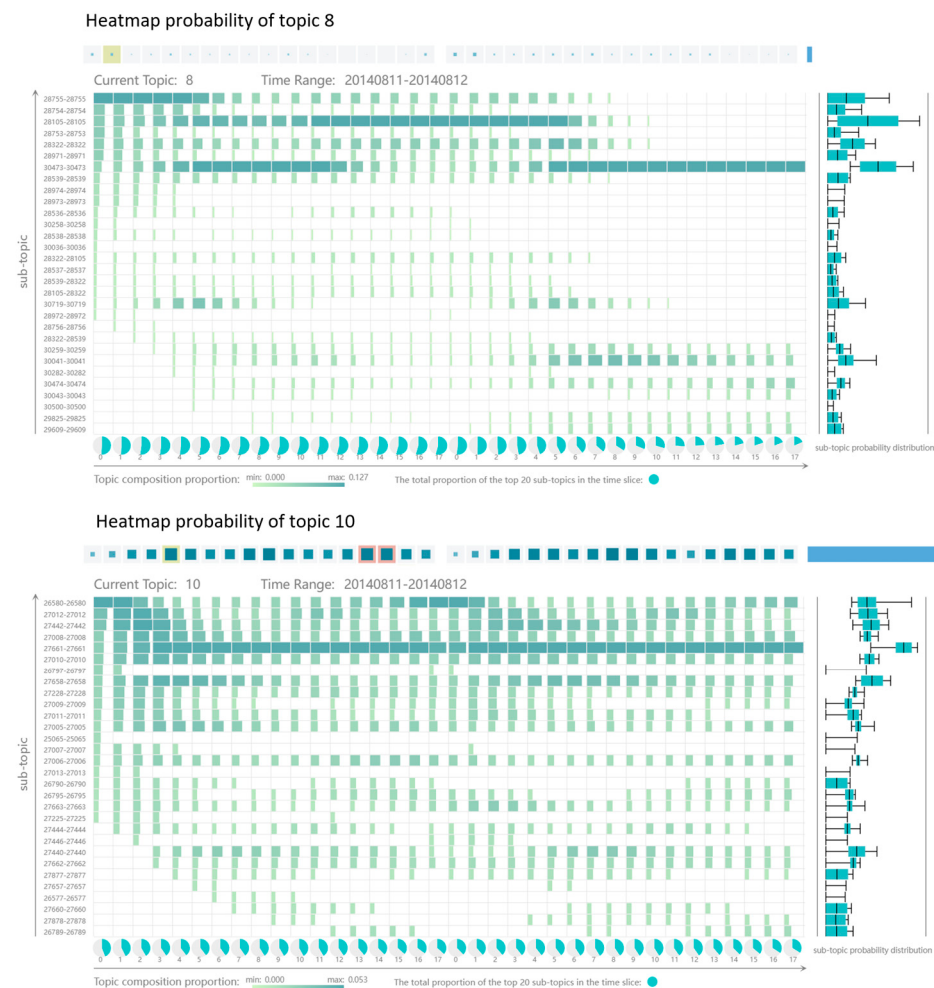


Figure 7. Different trajectory topics and their evolution exhibit significant differences. We compared the composition evolution of trajectory words, using the matrix heatmaps of topic words for trajectories 8 and 10. The figure captures the top 35 topic words that comprise the topic model and shows the composition proportion of the trajectory topic concerning the topic words (trajectory segment). The upper part of the topic word evolution matrix is the probability and topic entropy of the topic in different time slices. The pie chart on the lower side of the topic term evolution matrix gives the highest probability of 35 topic terms and the proportion of the probability sum of all topic terms. On the right side of the topic term evolution matrix is a box graph about the probability of the topic term, which reflects the distribution of the topic term over time.

5.2. Trajectory Topic Sequence Analysis

The trajectory sequence reflects the movement patterns within the trajectory topics. The trajectory sequence is generated through the improved MDP, incorporating the spatial information of trajectory words, and represents the most representative trajectory sequences within the trajectory topics. In computing trajectory topics, the trajectory sequences have been simplified by partitioning the map into grids. Concentrating trajectories within these grids facilitates the generation of trajectory sequences. This method transforms trajectory sequences into movements within a grid and between grids. We believe that movements within a grid are significant because, in the original data entries, a vehicle might not travel a great distance, and not in all records does a vehicle's movement span across grids. Figure 8 displays the representative trajectory topics generated through an enhanced MDP,

presenting an overview of the topics using an abstract map approach. Due to scenarios in the computational data where vehicles move within a single grid, sequences of movements within a single grid arise, as illustrated in Figure 8d,e. This reflects real situations present in the trajectory data. Additionally, trajectory sequences generated over time in topics 6 and 12 can supplement the trajectory direction information on the map, indicating the most probable movement patterns under that time's trajectory topic.

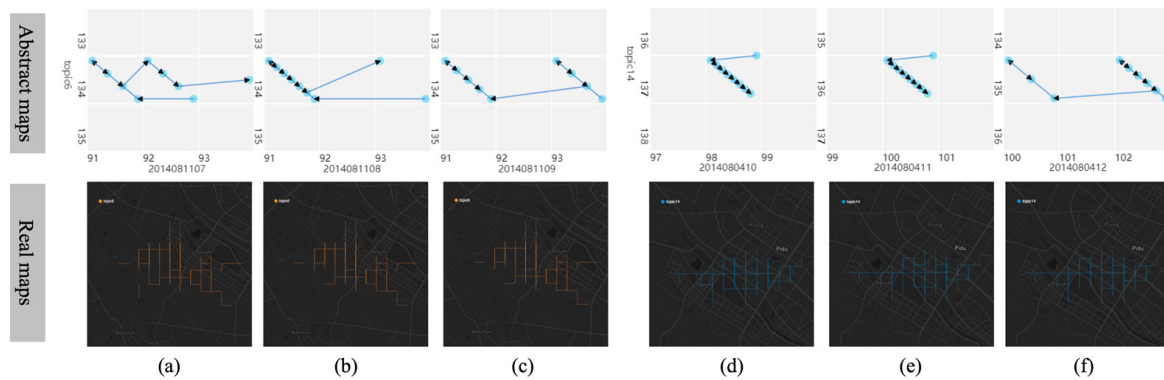


Figure 8. Analysis of trajectory topic sequence through abstract map and real map. The first row is the abstract map generated by improving the MDP, and the second row is the real map corresponding to the time and topic. (a–f) represent the combinations for abstract map and real map at different time slices.

6. Discussion

The core of the dynamic trajectory topic model is the evolution of trajectory topics. Compared to previous works on trajectory topic computation [2,9,11,14,18], this study introduces a method for the temporal evolution of topics. This method facilitates dynamic trajectory topic analysis by connecting the model parameters to the topic models of adjacent windows. It is more suited to real-world scenarios than matrix factorization-based trajectory topic analysis methods such as NMF.

The order of topic words is an essential component of the semantic structure in dynamic trajectory topics. Moreover, while the topic model is a probabilistic model regarding the composition of topic words, previous works did not consider the relationships between topic words, that is, the sequence of trajectory segments. In this study, an enhanced MDP, grounded on the probabilities of topic words, is introduced for generating representative trajectory sequences under specific trajectory topics. Our approach preserved the original data characteristics without imposing intricate feature constraints on the applied trajectory dataset. We hope this method can inspire similar research to focus on the inherent features of trajectory data and the continuity of trajectory information in the real world. However, there are certain limitations to our study. During the computation of dynamic trajectory topics, due to computational constraints inherent to the LDA model, a significant number of computational resources is required when fitting the LDA. Furthermore, there is a need to explore new trajectory point aggregation methods to compress the original trajectory point data for more efficient sequences.

7. Conclusions

A dynamic topic analysis approach for traffic trajectory data was proposed in response to the issue of existing trajectory visualization analysis methods overlooking the temporal evolution of trajectory topics. This facilitates the visual exploration and analysis of traffic trajectory topics and their evolution. Initially, the DTM is employed to characterize the dynamic topics of trajectories, seamlessly integrating both temporal and spatial information into the DTM. Subsequently, we propose an enhanced Markov decision process to generate topic trajectory sequences, thereby incorporating trajectory sequence information into the

trajectory topics. Following this, visual views designed based on parallel windows are introduced for exploring the evolution of trajectory topics. The objective is to convey information about dynamic trajectory topics through concise and intuitive visualization components. Ultimately, the efficacy of our methodology is illustrated through a case study focusing on the evolution of trajectory topics.

Since the trajectory dynamic topic model algorithm is based on a variational method to fit a posteriori distribution, there is room to enhance its computational efficiency. In the future, we plan to improve the efficiency of the dynamic trajectory topic model algorithm by adopting distributed learning algorithms, such as federated learning. Moreover, future research could explore the patterns and relationships of dynamic trajectory topics within and outside grids under various grid partitioning methods and investigate efficient and compact visualizations of spatiotemporal trajectory features.

Author Contributions: Conceptualization, H.C., Y.W. and W.Z.; software, H.T., J.L. (Jing Lei) and Z.W.; data preparation, G.W. and J.L. (Jing Liao); evaluation, H.C. and F.W.; Manuscript, H.C., G.W. and W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61872304.

Data Availability Statement: Data available in online website [29].

Conflicts of Interest: Authors Huaquan Tang and Jing Lei are employed by the company Mianyang Xincheng Engine Co., Ltd, Mianyang, Sichuan, China. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

Latent Dirichlet Allocation (LDA). Non-negative Matrix Factorization (NMF). Biterm Topic Model (BTM). dynamic topic model (DTM). Markov Decision Process (MDP).

References

1. Liu, L.; Zhang, H.; Liu, J.; Liu, S.; Chen, W.; Man, J. Visual exploration of urban functional zones based on augmented nonnegative tensor factorization. *J. Vis.* **2021**, *24*, 331–347. [[CrossRef](#)]
2. Liu, D.; Xu, P.; Ren, L. TPFlow: Progressive Partition and Multidimensional Pattern Extraction for Large-Scale Spatio-Temporal Data Analysis. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 1–11. [[CrossRef](#)] [[PubMed](#)]
3. Zhou, Z.; Meng, L.; Tang, C.; Zhao, Y.; Guo, Z.; Hu, M.; Chen, W. Visual Abstraction of Large Scale Geospatial Origin-Destination Movement Data. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 43–53. [[CrossRef](#)] [[PubMed](#)]
4. Deng, Z.; Weng, D.; Liang, Y.; Bao, J.; Zheng, Y.; Schreck, T.; Xu, M.; Wu, Y. Visual cascade analytics of large-scale spatiotemporal data. *IEEE Trans. Vis. Comput. Graph.* **2021**, *28*, 2486–2499. [[CrossRef](#)] [[PubMed](#)]
5. Andrienko, G.; Andrienko, N.; Fuchs, G.; Wood, J. Revealing Patterns and Trends of Mass Mobility Through Spatial and Temporal Abstraction of Origin-Destination Movement Data. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 2120–2136. [[CrossRef](#)] [[PubMed](#)]
6. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
7. Lee, D.; Seung, H.S. Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.* **2000**, *13*, 535–541.
8. Chu, D.; Sheets, D.A.; Zhao, Y.; Wu, Y.; Yang, J.; Zheng, M.; Chen, G. Visualizing Hidden Themes of Taxi Movement with Semantic Transformation. In Proceedings of the 2014 IEEE Pacific Visualization Symposium, Yokohama, Japan, 4–7 March 2014; pp. 137–144.
9. Liu, L.; Zhan, H.; Liu, J.; Man, J. Visual analysis of traffic data via spatio-temporal graphs and interactive topic modeling. *J. Vis.* **2019**, *22*, 141–160. [[CrossRef](#)]
10. Liu, H.; Jin, S.; Yan, Y.; Tao, Y.; Lin, H. Visual analytics of taxi trajectory data via topical sub-trajectories. *Vis. Inform.* **2019**, *3*, 140–149. [[CrossRef](#)]
11. Tao, Y.; Tang, Y. Progressive visual analysis of traffic data based on hierarchical topic refinement and detail analysis. *J. Vis.* **2023**, *26*, 367–384. [[CrossRef](#)]
12. Chen, C.; Liu, Q.; Wang, X.; Liao, C.; Zhang, D. semi-Traj2Graph Identifying Fine-Grained Driving Style With GPS Trajectory Data via Multi-Task Learning. *IEEE Trans. Big Data* **2021**, *8*, 1550–1565. [[CrossRef](#)]
13. Blei, D.M.; Lafferty, J.D. Dynamic topic models. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA, 25–29 June 2006; pp. 113–120.

14. Liao, L.; Wu, J.; Zou, F.; Pan, J.; Li, T. Trajectory topic modelling to characterize driving behaviors with GPS-based trajectory data. *J. Internet Technol.* **2018**, *19*, 815–824.
15. Huang, L.; Wen, Y.; Guo, W.; Zhu, X.; Zhou, C.; Zhang, F.; Zhu, M. Mobility pattern analysis of ship trajectories based on semantic transformation and topic model. *Ocean Eng.* **2020**, *201*, 107092. [[CrossRef](#)]
16. Wallach, H.M. Topic modeling: Beyond bag-of-words. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA, 25–29 June 2006.
17. Mohammadiha, N.; Smaragdis, P.; Panahandeh, G.; Doclo, S. A state-space approach to dynamic nonnegative matrix factorization. *IEEE Trans. Signal Process.* **2014**, *63*, 949–959. [[CrossRef](#)]
18. Yao, F.; Wang, Y. Tracking urban geo-topics based on dynamic topic model. *Comput. Environ. Urban Syst.* **2020**, *79*, 101419. [[CrossRef](#)]
19. Chen, W.; Huang, Z.; Wu, F.; Zhu, M.; Guan, H.; Maciejewski, R. VAUD: A visual analysis approach for exploring spatio-temporal urban data. *IEEE Trans. Vis. Comput. Graph.* **2017**, *24*, 2636–2648. [[CrossRef](#)]
20. Liao, C.; Chen, C.; Zhang, Z.; Xie, H. Understanding and visualizing passengers' travel behaviours: A device-free sensing way leveraging taxi trajectory data. *Pers. Ubiquitous Comput.* **2019**, *26*, 491–503. [[CrossRef](#)]
21. Havre, S.L.; Hetzler, E.G.; Nowell, L.T. ThemeRiver: Visualizing theme changes over time. In Proceedings of the IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings, Salt Lake City, UT, USA, 9–10 October 2000; pp. 115–123.
22. He, J.; Chen, C. Spatiotemporal analytics of topic trajectory. In Proceedings of the 9th International Symposium on Visual Information Communication and Interaction, Dallas, TX, USA, 24–26 September 2016; pp. 112–116.
23. Gao, X.; Liao, C.; Chen, C.; Li, R. Visual Exploration of Cycling Semantics with GPS Trajectory Data. *Appl. Sci.* **2023**, *13*, 2748. [[CrossRef](#)]
24. Al-Dohuki, S.; Wu, Y.; Kamw, F.; Yang, J.; Li, X.; Zhao, Y.; Ye, X.; Chen, W.; Ma, C.; Wang, F. Semantictraj: A new approach to interacting with massive taxi trajectories. *IEEE Trans. Vis. Comput. Graph.* **2016**, *23*, 11–20. [[CrossRef](#)]
25. Ali, F.; Kwak, D.; Khan, P.; El-Sappagh, S.; Ali, A.; Ullah, S.; Kim, K.H.; Kwak, K.-S. Transportation sentiment analysis using word embedding and ontology-based topic modeling. *Knowl.-Based Syst.* **2019**, *174*, 27–42. [[CrossRef](#)]
26. Yan, Y. Visual Analytics Based on Topic Models. Ph.D. Thesis, Zhejiang University, Hangzhou, China, 2019.
27. Zhou, Z.; Zhang, X.; Yang, Z.; Chen, Y.; Liu, Y.; Wen, J.; Chen, B.; Zhao, Y.; Chen, W. Visual Abstraction of Geographical Point Data with Spatial Autocorrelations. In Proceedings of the 2020 IEEE Conference on Visual Analytics Science and Technology (VAST), Salt Lake City, UT, USA, 25–30 October 2020; pp. 60–71.
28. Wang, H.; Ni, Y.; Sun, L.; Chen, Y.; Xu, T.; Chen, X.; Su, W.; Zhou, Z. Hierarchical visualization of geographical areal data with spatial attribute association. *Vis. Inform.* **2021**, *5*, 82–91. [[CrossRef](#)]
29. Chengdu Taxi GPS Data. Available online: <https://www.pkbigdata.com/common/zhzgbCmptDetails.html> (accessed on 21 January 2024).
30. Newman, D.; Lau, J.H.; Grieser, K.; Baldwin, T. Automatic evaluation of topic coherence. In Proceedings of the Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles, CA, USA, 2–4 June 2010; pp. 100–108.
31. Mimno, D.; Wallach, H.; Talley, E.; Leenders, M.; McCallum, A. Optimizing semantic coherence in topic models. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, 27–31 July 2011; pp. 262–272.
32. Aletras, N.; Stevenson, M. Evaluating topic coherence using distributional semantics. In Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)—Long Papers, Potsdam, Germany, 19–22 March 2013; pp. 13–22.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.