IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 3, 2021, accepted January 12, 2021, date of publication January 19, 2021, date of current version March 4, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052795

A Spatio-Temporal Co-Clustering Framework for Discovering Mobility Patterns: A Study of Manhattan Taxi Data

QIAN LIU[1,2], XINQI ZHENG[1,3], H. EUGENE STANLEY[2], FEI XIAO[4,5], AND WENCHAO LIU[4,5]

[1]School of Information Engineering, China University of Geosciences, Beijing 100083, China
[2]Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA
[3]Polytechnic Center for National Land Spatial Big-Data, Ministry of Natural Resources of the People's Republic of China, Beijing, Beijing 100036, China
[4]Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources of the People's Republic of China, Beijing 100036, China
[5]Information Center of Ministry of Natural Resources of the People's Republic of China, Beijing 100036, China

Corresponding author: Xinqi Zheng (zhengxq@cugb.edu.cn)

This work was supported in part by the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, under Grant KF-2020-05-063, and in part by the Fundamental Research Funds for the Central Universities under Grant 2652018085.

ABSTRACT Research on clustering spatio-temporal data to extract mobility patterns requires further development, as most existing studies do not simultaneously integrate data along both spatial dimensions and temporal dimensions but instead focus on only one dimension or separate the dimensions in analyses and applications, which could lead to discoveries that are not representative of the overall data or are dificult to interpret. To simultaneously reveal the spatial and temporal patterns of urban mobility datasets, we propose an analytical framework that is based on co-clustering and enables mobility behaviors to be distinguished in spatial and temporal dimensions. We use one month of taxi GPS data from the Manhattan area to explore spatio-temporal co-occurrence patterns. The spatial and temporal dimensions of taxi trip data were co-clustered by using the Bregman Block Average co-clustering algorithm with I-divergence (BBAC_I). We performed this process on weekdays and holidays and compared the mobility differences between these two periods. The experimental results demonstrated the effectiveness of this analytical framework, with which we can reveal the spatial patterns and their temporal dynamics as well as temporal patterns and their spatial dynamics in mobility data.

INDEX TERMS Mobility patterns, co-clustering, spatio-temporal co-occurrence, taxi trip.

I. INTRODUCTION

With rapid technological advancements in sensors and tracking facilities, large quantities of spatio-temporal data that register human mobility have become readily accessible and widely available[1]. These increasing volume datasets not only enable researchers to understand cities via data-driven technologies but also help planners integrate more information to improve decision-making in many application domains, such as geography [2], ecology [3], urban [4] and transportation systems [5]. However, the complexity of human movement data increases along with the complex dependence and interactions between temporal dimensions and spatial dimensions while exhibiting a distinct heterogeneity across space and time. For instance, in terms of a taxi trip, destinations of daily commuting are concentrated around workplaces (space) in the morning (time) and residential subdivisions (space) in the evening (time). As a result, the impact of human mobility on the urban transportation system or other corresponding urban systems is also heterogeneous across space and time. In this case, the need to introduce a data-driven approach that can retrieve interesting dynamic patterns and critical characteristics from a spatio-temporal concurrent perspective is urgently needed.

Spatio-temporal clustering is an important task in spatio-temporal pattern mining, which groups objects based on their spatial and temporal similarity and analyzes them at a higher level of abstraction. In urban activities analysis, spatio-temporal clustering is useful for monitoring and aggregating individual activities in specific regions and temporal features.

The associate editor coordinating the review of this manuscript and approving it for publication was Stefania Bonafoni.

Traditional spatio-temporal clustering methods focus on three aspects: (1) establish a temporal domain and measure the corresponding spatial distance between two objects [6], [7]; (2) underline the temporal order of objects at predefined locations [8]–[11]; and (3) separately measure spatial and temporal dimensions [12]–[15]. However, spatio-temporal clustering focuses on only one dimension (spatial or temporal) or separately analyzes the dimensions by clustering the spatial samples based on the similarities with the temporal features, or vice versa. Simple aggregation over a dimension (time and space) or separately analyzing each dimension could lead to discoveries that are not representative of the overall data and the loss of some spatio-temporal interaction information [16]. A clustering method that performs simultaneous clustering along both temporal and spatial dimensions is essential. Consequently, the goal of this article is to propose an analytical framework that simultaneously discovers global mobility patterns such as spatial distributions, temporal variations, and intuitive dynamic interaction patterns in the mobility data along both temporal and spatial dimensions.

Co-clustering is the task of simultaneously partitioning a data matrix that represents a joint probability distribution or co-occurrences between two random variables (both samples and features) [17], [18], which has been extensively applied for pattern analysis in many fields. The purpose of co-clustering is to identify clusters that exhibit similarity for a subset rather than an entire feature set, which has shown prominent performance improvement over traditional one-sided clustering algorithms [19], [20]. Therefore, the resulting co-clusters can reveal valuable implicit information in a dataset with two dimensions and overcome the loss of information on interaction features in two dimensions. For example, co-clustering gene expression data can be employed to identify the gene groups in specific experimental conditions (such as patients with AIDS) and simultaneously cluster documents and words to aid in understanding the semantics in a specific context. More importantly, co-clustering achieves excellent performance in partitioning sparse and high-dimensional data [21]. Recently, several studies also applied a feature-based co-clustering algorithm to analyze geo-referenced time series data in a variety of fields, such as phenological pattern extraction [3], disease hotspot detection [22], and identification of favorable conditions for virus outbreaks [23], [24]. In terms of analyzing human mobility, some researchers employed a feature-based co-clustering algorithm to extract features in OD flow data [16], [25], however, the objects of clustering are two spatial variables of the origin and destination. Therefore, co-clustering analysis is an extraordinarily appropriate technique for discovering hidden mobility patterns in spatio-temporal data, and it is also necessary to analyze it from the perspective of spatiotemporal co-occurrence.

In this article, we propose a framework that is based on a co-clustering algorithm to analyze spatio-temporal mobility patterns and their concise interaction characteristics from a spatio-temporal co-occurrence perspective. Experiments are performed using the taxi trip time series of NYC on both weekdays and weekends to demonstrate the effectiveness and practicality of the proposed framework. First, using one-month taxi GPS datasets for the Manhattan area of New York City (NYC), we reorder the mobility matrix that emerges with urban regions and the time series of a day based on the co-clustering method and obtains a checkerboard-like spatio-temporal co-clustering pattern. Second, we apply K-means to refine the checkerboard-like clustering pattern into non-checkerboard spatio-temporal co-clusters and identify the unique spatial and temporal patterns. Last, the spatial patterns with their temporal dynamics and the temporal mobility patterns with the corresponding spatial distributions are visualized to gain region- and timestamp-specific insights. To better understand the travel behavior and urban dynamic characteristics, we split the dataset into two groups in terms of periods as weekdays and holidays and perform this process by comparing the trip patterns between weekdays and holidays. The main contribution of this article is to propose an analytical framework for simultaneously revealing spatial and temporal patterns of transportation/mobility data. The adopted statistical techniques can adequately interpret latent spatio-temporal interactions and complex dependence in large amounts of spatio-temporal mobility data. It is noted that this article does not intend to develop a new clustering methodology for the employed methods. The methods that we apply for co-clustering and K-means are existing methods because our focus is on gaining new empirical insight by exploring mobility data rather than contributing to these data mining techniques.

The remainder of this article is organized as follows: Section II reviews relevant studies of spatial and temporal analysis of taxi trips and spatio-temporal clustering methods, in particular, the co-clustering algorithm. Section III describes the detailed process of checkerboard and non-checkerboard co-clustering methods that are applied to spatio-temporal data. Section IV performs experiments on taxi trip data from NYC on both weekdays and weekends to demonstrate the effectiveness and practicality of the proposed framework. Section V provides the conclusions of this article.

II. LITERATURE REVIEW

A. SPATIAL AND TEMPORAL ANALYSIS OF TAXI TRIP

The taxi trip is a significant mode of transportation in urban areas due to its flexible door-to-door service and 24-7 operation [26]. With information about when and where a customer is picked up or dropped off by a taxi, meaningful dynamic patterns of a city can be obtained by data mining approaches and models [27], [28].

Previous research has focused on analyzing the mobility patterns of the taxi trip. Some researches focused on the spatial content of the taxi trip. For example, considering the differences between a long distance and a short distance in the taxi GPS trajectory data of an intracity, [29] built network models and performed community analysis to reveal a two-level hierarchical polycentric structure of Shanghai

with a viewpoint of spatial interactions represented by taxi trips. Reference [30] proposed a matrix factorization method that is based on an analytical framework to detect taxicabs' operation patterns in space by analyzing their continuous digital traces. Conversely, some studies focused on temporal information about the taxi trip. For instance, by merging GPS records of taxi trips and historical weather records in NYC, [31] introduced a classification and regression tree methodology to analyze the temporal patterns of the descriptive statistics of travel time (such as average travel time, standard deviation, and coefficient of variation) and discovered that the peak periods exhibit inter period heterogeneity in terms of the average travel time and standard deviation, whereas the coefficient of variation exhibited more consistent patterns among the days. However, pattern extraction in these studies is imperfect because only the single features of time or space are considered.

To extract the more comprehensive information about mobility patterns, additional studies considered both spatial features and temporal features. Reference [32] investigated intra-urban human mobility by analyzing the temporal and spatial distributions of trip distances and trip directions and proposed a model that integrates both the geographical heterogeneity and distance decay effect to interpret these patterns. Reference [33] developed an effective model to explore the spatio-temporal characteristics of individuals' daily activities and determine the temporal variations of activities by showing a strong periodicity of activity intensity in 1 week. Notably, all these studies analyzed spatial and temporal characteristics separately. In addition, in recent years, there emerged a series of effective matrix factorization methods for detecting underlying spatio-temporal patterns and exploring the high-level interactions among different spatial, temporal, and other attributes in mobility data [16]. Matrix factorization is a local perception of the research object, which extracts and explains local features from multiple dimensions (such as temporal, spatial, and other attributes) [34]. The latent interactions between different dimensions are represented by the combined probability across different patterns in each dimension, and the mixed probability distribution makes the result unintuitive and difficult to interpret. Therefore, it is necessary to provide a more concise way to efficiently present the interaction characteristic among spatial and temporal patterns.

B. SPATIO-TEMPORAL CLUSTERING AND CO-CLUSTERING ALGORITHM

Spatio-temporal clustering is an important technology of spatio-temporal data mining that aims to extract a series of clusters based on both spaces and timestamps [35], [36]. Spatio-temporal clustering is also beneficial for investigating the distributions of geographical phenomena and detecting the spatio-temporal characteristics of clusters [37]. In recent years, spatio-temporal clustering has been extensively applied in many research fields, including socioeconomic analysis, climate change analysis, epidemic detection, crime behavior prediction, and traffic dynamic analysis [18], [38]–[41]. In terms of these different applications, clustering methods are primarily dependent on the specific characteristics of five types of spatio-temporal datasets: spatio-temporal events, geo-referenced variables, geo-referenced time series, moving objects, and trajectories [35]. In this study, we focus on clustering geo-referenced time series, which record time-changing values of one or more observed attributes at fixed locations and consistent time intervals [42].

Co-clustering enables us to directly aggregate objects based on their similarities with respect to observed nonspatial attribute values at different time stamps and regions. Since it was first proposed by [17], the importance of the concept of co-clustering has increased and has been extensively employed for exploratory analysis in many fields. Reference [43] proposed a spectral bi-clustering algorithm that simultaneously clusters genes and conditions of expression data and discovers distinctive ''checkerboard'' structures in eigenvectors that correspond to a characteristic expression pattern across genes or conditions. Reference [44] proposed information-theoretic co-clustering (ITCC), which uses the I-divergence matrix to minimize the mutual information loss between the original random variables and the clustered random variables and discovered groups of interrelated words and documents by employing text mining. Reference [45] proposed the Minimum Sum-Squared Residue Co-clustering (MSSRCC) to minimize two different objective functions based on two different squared residue measures. Reference [19] introduced the Bregman Block Average co-clustering (BBAC) algorithm with I-divergence (BBAC_I) to measure the approximation error using a large class of loss functions named Bregman I-divergences. The BBAC algorithm also enables the use of many other metrics(such as the squared Euclidean distance and KL-divergence) as loss functions to optimize the co-clusters. Among these metrics, the superiority of the BBAC_I for simultaneously analyzing two-dimension features (e.g., word-document) has been suggested. Some researchers recently applied the BBAC_I into the phenology field and successfully identified the spring phenology patterns along locations and years.

Even though previous studies have analyzed the mobility patterns of taxi trips, a framework that simultaneously examines the dynamic mobility patterns along both temporal dimensions and spatial dimensions is lacking. In this article, we introduce a co-clustering algorithm (BBAC_I) to extract similar spatio-temporal groups and their interaction characteristics of taxi trips in NYC from a spatio-temporal co-occurrence perspective. Although the co-clustering method is not new, this article focuses on its application to a spatio-temporal mobility dataset and reveals valuable insight into spatio-temporal-mobility interactions.

III. METHODOLOGY

In our research, the presented approach includes two main parts to achieve simultaneous retrieval of spatial and temporal patterns for human activities that are obtained from taxi trip

data. The geo-referenced time series data is first divided into checkerboard-like co-clusters that are composed of spatial and temporal dimensions using a co-clustering algorithm, and then K-means is applied to regroup the regular spatio-temporal co-clusters and refine the spatio-temporal interaction patterns.

A. THE BREGMAN BLOCK AVERAGE CO-CLUSTERING ALGORITHM WITH I-DIVERGENCE

Co-clustering is an important pattern analysis tool that simultaneously clusters rows and columns of a two-dimensional data matrix. Retrieved checkerboard co-clusters reveal the intrinsic associations between rows and columns[46]. A popular means of co-clustering uses information-theoretic cost functions, which consider co-clustering as a lossy data compression problem and attempt to minimize the loss in mutual information between the original data matrix and the co-clustered matrix [3], [44], [47] (or retain as much information as possible about the original data matrix [19]). Based on the description of the co-clustering algorithm in the literature review, the BBAC algorithm enables the use of multiple metrics (such as the squared Euclidean distance, KL-divergence, and I-divergence) as cost functions to optimize the co-clusters. The superiority of I-divergence over other metrics has been empirically shown [19].

In our research, the BBAC_I algorithm is employed to capture spatio-temporal-mobility interactions and simultaneously group data elements based on their similarity along regions and timestamps of the original data matrix $O(R, T)$ (which is also the spatio-temporal co-occurrence matrix) into the co-clustered matrix $\mathrm{C}(\widehat{R}, \widehat{T})$. The spatio-temporal co-occurrence matrix can be regarded as the co-occurrence distribution between two random variables: the regions (R) and the timestamps (T). In this context, the rows of the co-occurrence matrix refer to the region sets $\{r1, r2, \ldots, rm\}$ that represent the fixed spatial area, and the columns represent the regular intervals $\{t1, t2, \ldots, tn\}$ that belong to one day. The elements of the co-occurrence matrix are the average travel volumes for each region and time period. We are devoted to obtaining a $k \times l$ partitional co-clustering matrix $\mathrm{C}(\widehat{R}, \widehat{T})$, which quantizes R into k region-cluster sets $\widehat{R}$: $\{\widehat{r1}, \widehat{r2}, \ldots, \widehat{rk}\}$ ($\mathrm{k} \leq \mathrm{m}$) and T into l timestamp-cluster sets $\widehat{T}$: $\{\widehat{t1}, \widehat{t2}, \ldots, \widehat{tl}\}$($\mathrm{l} \leq \mathrm{n}$). In addition, the mutual information between the spatial variable R and the temporal variable T is denoted as $I(\mathrm{R}; \mathrm{T})$, similarly, and $I(\widehat{R}; \widehat{T})$ is the mutual information between the spatial variable $\widehat{R}$ and the temporal variable $\widehat{\mathrm{T}}$. Therefore, the BBAC_I algorithm is applied to achieve an optimization problem, where the loss in mutual information $I(R; T) - I(\widehat{R}; \widehat{T})$ (note that $I(R; T) \leq I(\widehat{R}; \widehat{T})$) between the original data matrix $O(R, T)$ and the optimum co-clustered matrix $\mathrm{C}(\widehat{R}, \widehat{T})$ is minimized.

The detailed optimization process is shown in Table 1. The original data matrix contains the average travel volumes for m regions (R) and n timestamps (T). The number of region-clusters $\widehat{R}$ and temporal-clusters $\widehat{T}$ is predefined by the user as inputs, which represent the interaction degree between the spatial dimensions of the mobility patterns and the temporal dimensions of the mobility patterns. The first step of the BBAC_I algorithm is to randomly initialize the mapping of regions to region-clusters and the mapping of timestamps to timestamp-clusters, which yields the co-clustered matrix $\mathrm{C}(\widehat{R}, \widehat{T})$. In the second step, the loss in mutual information between the original matrix and the co-clustered matrix (which is composed of the region-clusters and timestamp-clusters) is calculated by I-divergence metrics, where $\mathrm{D}(\cdot \parallel \cdot)$ denotes the I-divergence between two matrixes. In the third step, an iteration process is initiated to separately update the mapping from regions to region-clusters and the timestamps to timestamp-clusters by minimizing the loss function listed in this step. Last, the loss of mutual information is recalculated before and after performing the co-clustering. If the change in the loss of mutual information is smaller than a predefined threshold, the optimal co-clustering result $\mathrm{C}(\widehat{R}, \widehat{T})$ is obtained and grouped by the optimal mappings $\widehat{R}$ and $\widehat{\mathrm{T}}$ in step 3; otherwise, we return to the iteration process. The rows and columns of the original spatio-temporal co-occurrence matrix are regrouped into checkerboard-like co-clusters, where regions and timestamps that belong to the same region-cluster or timestamp-cluster are concurrently arranged.

TABLE 1. Algorithm table of the BBAC_I for discovering mobility patterns from spatio-temporal co-occurrence data.

Input: $O(R, T)$: original data matrix, k: number of region-clusters, l: number of timestamp-clusters

Steps:

1. Initialization: randomly initialize mapping from m regions to k region-clusters and n timestamps to l timestamp-clusters;
2. Calculate the loss in mutual information
$$I(R;T) - I(\hat{R};\hat{T}) = \mathrm{D}(O(R,T) \parallel \mathrm{C}(\hat{R},\hat{T}))$$
3. Iteration:
   - 3.1 Update mapping from regions to new region-clusters
   $$\hat{R} = \arg\min_{\hat{r} \in \{1,\ldots,k\}} D(O(\mathrm{R},T) \parallel \mathrm{C}(\hat{R},\hat{T}))$$
   - 3.2 Update mapping from regions to new region-clusters
   $$\hat{T} = \arg\min_{\hat{t} \in \{1,\ldots,l\}} D(O(\mathrm{R},T) \parallel \mathrm{C}(\hat{R},\hat{T}))$$
4. Convergence: once the loss in mutual information is minimal, the update of the $\hat{R}$ and $\hat{T}$ matrices stops, and the optimal co-clustered results are obtained; otherwise, return to Step 3.

Output: $\mathrm{C}(\hat{R},\hat{T})$: Optimized spatio-temporal co-cluster

B. EXTRACTING NON-CHECKERBOARD CO-CLUSTERS

Based on the BBAC_I algorithm, a distinctive checkerboard clustering pattern can be obtained. The elements in the same block are similar and represent the aggregation of regions and time period with similar mobility patterns. However, the similarity among the values in the different co-clusters (blocks) may be attributed to the assignment of full rows and columns to region-clusters and timestamp-clusters in the BBAC_I algorithm. For example, the travel volume is scarce in most regions at midnight, which causes similar values of co-clusters in the midnight patterns. A more efficient way to manipulate the regular co-clusters for pattern extraction is therefore needed.

To tackle the limitation of checkerboard clustering, we combine the well-known K-means clustering to regroup the regular spatio-temporal co-clusters and refine the spatio-temporal interaction patterns, which is based on the principle of ''divide and group''. Divide and group is one of the fundamental principles in exploratory data analysis [48], which applies when the overall behavior exhibited in a selected dataset (called reference set) can be represented by aggregating behaviors in the subsets of the reference set. The ''divide'' and ''group'' are two interrelated analysis processes, which can promote pattern discovery. The process of dividing a dataset according to the selected criteria produces several subsets of the dataset, each of these subsets contains data objects grouped by the selected criteria, and hence the grouping process [49]. Data attributes are often used as criteria, and clustering is the most commonly used method to divide and group data. Therefore, the principle of ''divide and group'' is suitable for our research to improve the flexibility of the checkerboard clustering, and achieve the interactive classification of mobility dataset [50].

At this part, we use K-means to group the checkerboard-like co-clustering results into k axis-parallel non-checkerboard co-clusters, as noted by [3]. We calculate the mean and standard deviation of each block in the checkerboard-like co-clusters and generate a $(k \times l) \times 2$ input matrix (based on the co-cluster numbers in Section III.A) for K-means clustering, because the former provides a representative value of elements in each checkerboard-like co-cluster and the latter considers the presence of possible outliers within each checkerboard-like co-cluster, which present the overall distribution and local deviation of the original data well. The optimal number of the K-means' clusters are evaluated and determined using two indicators: the Silhouette coefficient and the sum of the squared errors (SSE) [51]. The Silhouette coefficient can be understood as an index for describing the sharpness of the Silhouette of each cluster after clustering; the value range of this index is -1 to 1 to 1. The larger the average Silhouette coefficient of all samples is, the better the clustering effect. In terms of the SSE, it is usually combined with the elbow method to obtain the optimal cluster number (K). With an increase in K, the aggregation degree of each cluster will gradually increase, and the SSE will naturally decrease. The best clustering is achieved when the decline of the SSE curve suddenly decelerates, and the K value that corresponds to the elbow-shaped split point is the best clustering number. Therefore, the clustering value with the larger Silhouette coefficient and proper ''elbow'' split point is the best choice.

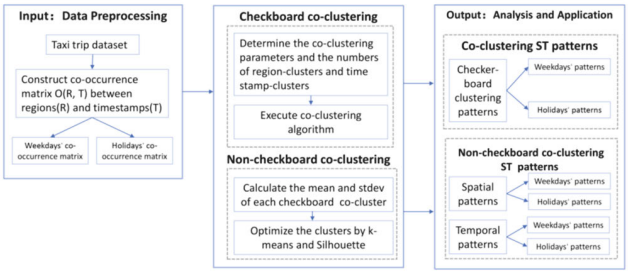


FIGURE 1. The framework of the co-clustering spatio-temporal analysis.

IV. EXPERIMENTS AND DISCUSSION

In this section, we introduce the BBAC_I algorithm to analyze the spatio-temporal patterns of a taxi trip in Manhattan. To better understand the trip characteristics, an empirical analysis is conducted on both weekdays and holidays. FIGURE 1 displays the research framework proposed in this article. The data preprocessing is illustrated in Section IV.A; we construct two spatio-temporal co-occurrence matrixes at 2 h intervals for weekdays and holidays. In Section IV.B, the co-clustering parameters and the size of the cluster before the co-clustering analysis is determined. Section IV.C describes the checkerboard co-clustering patterns and corresponding spatial distribution of the checkerboard co-cluster for weekdays and holidays. Section IV.D reveals the human mobility spatio-temporal patterns of weekdays and holidays using non-checkerboard co-clustering, which simultaneously examines both the spatial patterns and their temporal dynamics in taxi trips as well as the temporal patterns and their spatial dynamics. Section IV.E compares our method with several traditional clustering methods and discuss the limitation of our research, as well as future work.

A. STUDY AREA AND DATASETS

Taxis have significant roles in public transport systems in many metropolitan areas. Due to their popularity and importance, numerous studies utilize the mobility data of taxis to explore the metropolitan rhythm and pulse, such as traffic flow and human mobility patterns. According to a report published by the NYC Taxi and Limousine Commission (NYCTLC), more than 17,000 licensed taxis (yellow and green taxis) generated approximately 343,000 trips daily in 2017 [52]. The enriched mobility data provided the opportunity to extract and analyze spatio-temporal mobility patterns using the co-clustering methodology. In this article, we investigate taxi rides in NYC—one of the most densely populated international metropolises in the world. In New

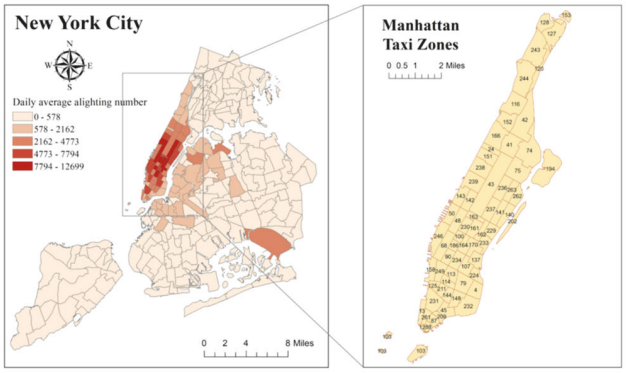


FIGURE 2. Study area and spatial distribution of the average daily drop-off number of taxi trips in November 2017 across the NYC area.

York City, there are several types of licensed vehicles, which include medallion taxis (''yellow cabs''); street hail liveries (''boro or green taxis''); black cars, liveries, and luxury limousines (''FHVs''); commuter vans; paratransit vehicles; and wheelchair accessible vehicles (''WAVs''). Among them, the services of yellow and green cabs are permitted to pick up passengers via street hails, thus offering a great footprint of human activity. Furthermore, yellow cabs are concentrated in Manhattan and two main airports, and green cabs are allowed to service the area above E. 96th St. and W. 110th St. in Manhattan and anywhere in the other boroughs. Therefore, combining these two datasets can acquire good coverage of the entire city, which is also research in [53]. The taxi trip datasets (of both yellow and green taxis), which were downloaded from the official website of the NYC Taxi and Limousine Commission (https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page), were collected from November 1, 2017, to November 31, 2017, where there are approximately 10 million distinct trip records totally. Each trip record contained spatial information (pick-up location and drop-off location) and temporal information (pick-up timestamp and drop-off timestamp) for both pick-up activities and drop-off activities; the trip distance, fare, and the number of passengers were also provided.

For this case, we select the drop-off trips to construct a co-occurrence matrix. In other words, we analyze taxi passengers' destination choice activities. The distribution of daily average arrival data for the NYC area in November 2017 is visualized, as shown in FIGURE 2(a). The data indicate that the destinations of taxi trips are concentrated in Manhattan and its surrounding areas. In this article, we focus our analysis on the Manhattan district, with the exception of the three island regions (Governor's Island, Ellis Island, and Liberty Island), for which the daily average drop-off volume is zero. The study area is shown in FIGURE 2(b), and each region is labeled by the location ID (a total of 66 regions; the geographic information of regions is provided by the TLC). To investigate the temporal characteristics of trips, we split the initial dataset into two parts in terms of time periods, namely, weekdays and holidays, which contribute to a better understanding of the dynamic patterns with different human mobility. For the weekday trips, we counted the typical working day (Tuesday to Thursday) in November 2017 (excluding Veterans Day, Thanksgiving, and the Wednesday before Thanksgiving), which total 12 days. Holiday trips encompass 9 days, including Saturday, Sunday, and Thanksgiving Day. For each temporal part, an interval of 2 h is selected to divide a full day of operations (24 h) into 12 time periods[54, 55], and counted the number of taxis that drop-off in each region for each time period (average alighting data for 2 h intervals in one day). Therefore, the dynamic drop-off matrix $O(R, T)$, which can be regarded as a spatio-temporal co-occurrence extent between the regions (R = 66 rows) and timestamps (T = 12 columns), was obtained.

B. DETERMINATION OF THE CLUSTER SIZE

When applying co-clustering, the first issue to consider is the cluster size, which represents the interaction degree between the spatial dimension and the temporal dimension. The determination of the clustering number requires comprehensive consideration of the characteristics of the data, the purpose of clustering, and the validity of the clustering effect. If the cluster size is large, the risk of overfilling may increase. If the cluster size is too small, the clusters in the latent space are not adequately depicted and determined [28], which may disregard the complex spatio-temporal interdependence of the trip data. In this case, the number of timestamp-clusters was set to 4 after testing values from 4 to 8 in the initialization process of the BBAC_I algorithm because the co-clustered results always returned this number of timestamp-clusters, which means the loss function of BBAC_I reached minimum with this value. Similar to the temporal characteristics of previous studies [34], [56], the divided timestamp-clusters can be used to characterize four time periods: the early peak period, evening peak period, off-peak period, and late-night period. The specific intervals for these periods are determined by the results of temporal clustering. For the spatial dimension, we test the values of the region-clusters (k) from 4 to 15 a total of 5 times to minimize the values of the loss function (objective function) in the BBAC_I algorithm, which identifies the optimal number of region-clusters. FIGURE 3 shows the change in the objective function with k for weekdays and holidays, respectively. As shown in FIGURE 3(a), in the interval [4], [8], the objective function decreases with an increase in k; in the interval [8], [15], the objective function becomes relatively stable. For weekdays, when k = 8 is selected, the objective function reaches the minimum value of the interval [3], [18], which has excellent interpretability of the experiment results. For holidays, the number of region-clusters is 6. Therefore, we conduct a co-clustering analysis with the cluster size 4 × 8 (region-clusters × timestamp-clusters) for weekdays and a co-clustering analysis with the cluster size 4 × 6.

Similar to other cluster algorithms, three predefined parameters exist in the BBAC_I algorithm: the times of random mapping initializations, the number of iterations, and the threshold for convergence with the loss of mutual

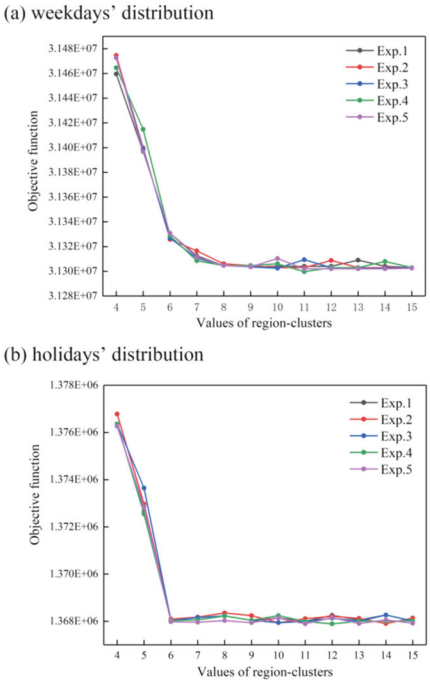


FIGURE 3. Change in objective function with the values of region-clusters. The x-axis represents the test values of the region-clusters from 4 to 15. The different colored curves represent the results of five experiments, which show the change in the values of the loss function. (a) weekday distribution; (b) holiday distribution.

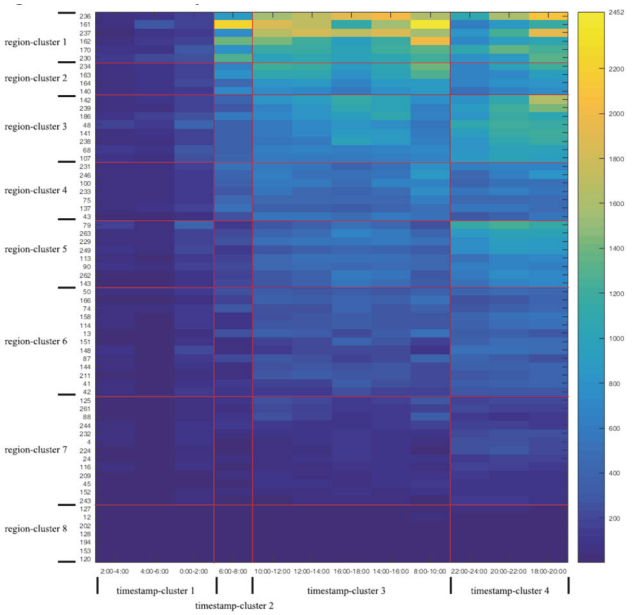


FIGURE 4. Weekdays' checkerboard co-clusters intersected by eight region-clusters and four timestamp-clusters. Each region-cluster and timestamp-cluster are partitioned by a thin red line. The color of the blocks represents the drop-off volume in the corresponding region and timestamp, which vary from dark blue, which represents Very Low volume, to light yellow, which represents high volume.

information. In this study, to determine the global optimal, we ran the BBAC_I algorithm 200 times with different random initial values. The convergence criteria were set to the threshold of the changes in the loss of mutual information, which was less than 10^{-6}. The maximum number of iterations was set to 2000 to guarantee convergence stability and the quality of the optimal co-clustering results.

C. CHECKERBOARD CO-CLUSTERING SPATIO-TEMPORAL PATTERNS

In this section, we applied the BBAC_I algorithm to map the weekday and holiday co-occurrence matrixes into 4×8 and 4×6 checkerboard co-clusters. FIGURE 4 and FIGURE 5 use heatmaps to exhibit the checkerboard co-clustering patterns for weekdays and the checkerboard co-clustering patterns for holidays, respectively, which intuitively aids in understanding specific travel characteristics among different spatio-temporal distributions. The x-axes represent the reordered 12 timestamps that were grouped into four timestamp-clusters. The y-axes represent the reordered 66 regions, which were mapped into eight region-clusters for weekdays and six region-clusters for holidays. The timestamp-clusters are reordered from the left of x-axes to the right of x-axes based on the average drop-off volume in each timestamp-cluster over the entire study area from less to more, and the timestamp-column in each timestamp-cluster is arranged with an increase in the drop-off volume from left to right. From the bottom to the top of the y-axes, the region-clusters are arranged in the order from region-cluster 8 to region-cluster 1 (or region-cluster 6 to region-cluster 1) with an increase in the average drop-off volumes of region-clusters for the whole day on weekdays (or holidays). The region-row in each region-cluster is arranged with the increase in drop-off volume from bottom to top. Based on this overall arrangement, the cells in each checkerboard are also arranged with an increase in the drop-off volume from left-bottom to right-top.

FIGURE 4 describes the checkerboard spatio-temporal co-clustering patterns for weekdays. We observe distinct commuting temporal patterns: timestamp-cluster 1 (0:00-6:00 for the late-night period), timestamp-cluster 2 (6:00-8:00 shows the off-peak period for most regions), timestamp-cluster 3 (8:00-18:00 for the daytime period; the early peak appears from 8:00 to 10:00, which is the rightmost column in this timestamp-cluster) and timestamp-cluster 4 (18:00-24:00 for evening periods; the evening peak appears from 18:00 to 20:00, which is the rightmost column in the heatmap). The rhythms of the taxi activities considerably vary among different timestamp-clusters. Compared with other time-clusters, the liveness of all regions, with the exception of region-cluster 8, is the highest in timestamp-cluster 4 (the period from 18:00 to 24:00), which indicates that taxi trips in the whole Manhattan area are most active in the evening.

The checkerboard spatio-temporal co-clustering patterns on holidays are shown in FIGURE 5. Similar to weekdays, the timestamps are partitioned as timestamp-cluster 1 (4:00-8:00 for the late-night period), timestamp-cluster 2 (2:00-4:00 and 8:00-10:00 show the off-peak periods), timestamp-cluster 3

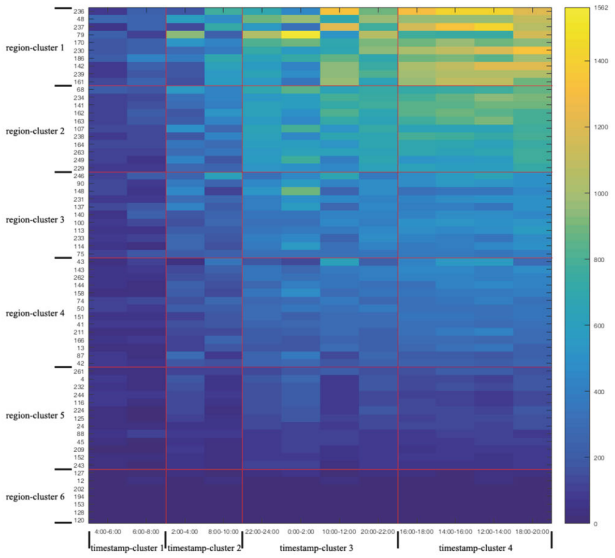


FIGURE 5. Holidays' checkerboard co-clusters intersected by six region-clusters and four timestamp-clusters, which are indicated by the thin red line.

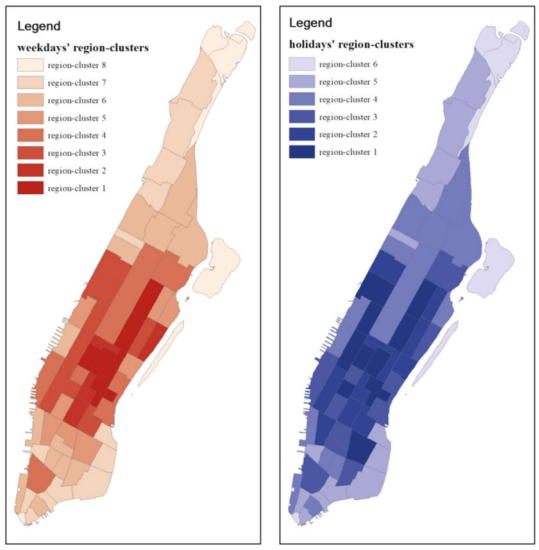


FIGURE 6. Spatial distribution of region-clusters in the checkerboard co-clustering patterns. Eight region-clusters exist for weekdays, and six region-clusters exist for holidays, which are labeled by the sequence of each region-cluster. The activities of each area are represented by color, where light yellow indicates a Very Low drop-off volume and red indicates a high drop-off volume.

(20:00-02:00 and 10:00-12:00) and timestamp-cluster 4 (12:00-20:00). The temporal distribution of taxi activities on holidays is more evenly distributed compared to the early and evening peak commuting patterns on weekdays, and the peak time was concentrated from 12:00 am to 20:00 pm. As we usually observe, the temporal distributions are delayed by more than two hours on weekdays. Unlike the smooth changes in the drop-off volume from region to region on weekdays, some conspicuous distributions exist on holidays, such as the block that represents the drop-off volume in the 79^{th} region from 12 am to 2 am. This finding suggests that taxi activities in different regions on holidays are more scattered or random.

To better obtain spatial visualization results, we exhibit the spatial distribution of region-clusters in FIGURE 6. Although the co-clustering process does not consider spatial location information, strong spatial agglomerating characteristics exist for both weekdays and holidays. The spatial distribution of taxi activities is concentrated in the central area of Midtown Manhattan and the regions adjacent to Central Park on the Upper East Side. Compared with holidays, the spatial distribution of weekday' trips is more concentrated in the core area of the CBD in Midtown, whereas most drop-off trips on holidays emerge in the regions of Upper West Side and Downtown (which are represented by region-cluster 1 in FIGURE 5).

D. NON-CHECKERBOARD SPATIO-TEMPORAL CO-CLUSTERING PATTERNS

Based on the co-clustering method, we identify the checkerboard co-clusters within the similar drop-off volume along the spatial dimension (regions) and temporal (timestamps) dimension. However, some co-clusters have similar drop-off volumes, for instance, the co-cluster in region-cluster 5/timestamp-cluster 3 and the co-cluster in region-cluster 4/timestamp-cluster 4 in FIGURE 4. To better extract accurate co-clustering spatio-temporal patterns, we regroup the checkerboard co-clusters into axis-parallel non-checkerboard co-clusters using K-means clustering. The mean and standard deviation of each checkerboard co-cluster were employed as the attribute dimensions of the samples in K-means. Therefore, a total of 32 samples for weekdays and 24 samples for holidays exist; the corresponding values of the mean and standard deviation are shown in FIGURE 7.

Because the x-axis of the curve in FIGURE 7 is sorted by the column sequence of the co-cluster matrixes, it can be seen that the curves of the mean and standard deviation are divided into four gradients, which belong to four timestamp-clusters in the order from timestamp-cluster 1 to timestamp-cluster 4. As shown in FIGURE 7(b), the overall mean of the drop-off volume sustained growth from timestamp-cluster 1 to timestamp-cluster 4 on holidays. In each timestamp-cluster, the mean value of the first row (corresponds to the cluster sequence of 1, 7, 13, 19, which belongs to region-cluster 1) in the co-cluster matrixes is the largest, and the mean value of the sixth row in the co-cluster matrixes is the smallest (corresponds to the cluster sequence of 6, 12, 18, 24, which belongs to region-cluster 6). For the weekday distribution, except for the fluctuation in timestamp-cluster 4, the mean distributions are similar to the holiday distribution: the mean value of the first row (corresponds to the cluster sequence of 1, 9, 17, 25, which belongs to region-cluster 1) in the co-cluster matrix is the largest, and the mean value of the eighth row in the co-cluster matrix is the smallest (corresponds to the cluster sequence of 8, 16, 24, 32, which

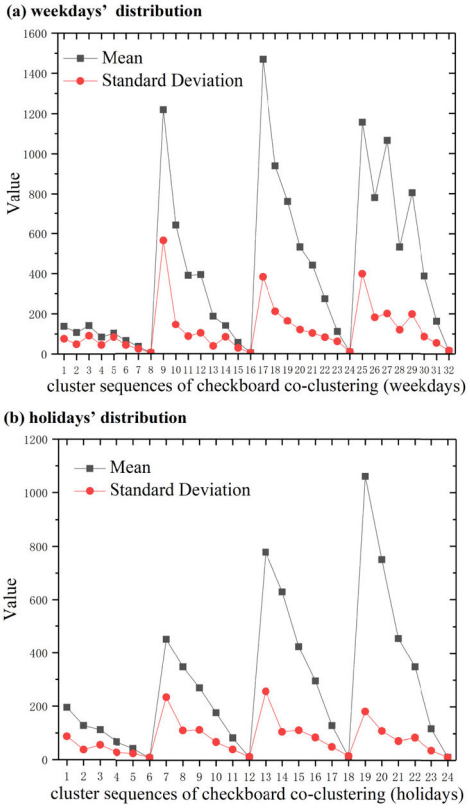


FIGURE 7. Mean and standard deviation of each checkerboard co-cluster. The x-axis represents the cluster sequences, which are arranged based on column-wise filling in checkerboard co-cluster matrixes for weekdays and holidays in FIGURE 4 and FIGURE 5. The black square curve represents the mean, and the red dotted curve indicates the standard deviation; their corresponding values are shown on the y-axis. (a) weekday distribution; (b) holiday distribution.

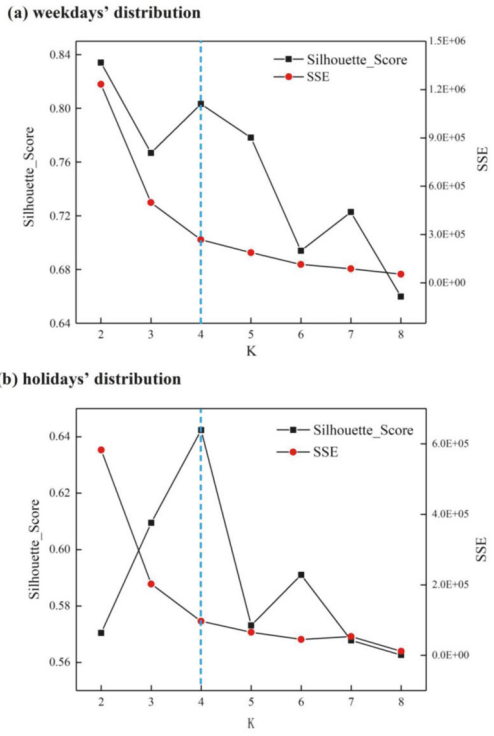


FIGURE 8. Changes in Silhouette and SSE results with respect to different K values. The black square curve represents the average Silhouette coefficient of all samples, and the red dot curve indicates the SSE. Their corresponding values are shown on the left side and right side, respectively, of the y-axis. (a) weekday distribution; (b) holiday distribution.

belongs to region-cluster 8). In addition, some peaks are observed in the distributions of the standard deviation for both weekdays and holidays, and the corresponding cluster sequence belongs to region-cluster 1, which indicates that regions in region-cluster 1 present a relatively higher disparity of drop-off volume than other region-clusters. Based on the values of the mean and standard deviation, we conduct K-means clustering to test the value of K from 3 to 8 and use the Silhouette coefficient and SSE to select the optimal number of non-checkerboard co-clusters. FIGURE 8 shows the changes in the Silhouette and SSE results with respect to different K values for weekdays and holidays. The SSE decreases from $K = 2$ and ushered in a turning point (elbow-shaped split point) with $K = 4$ in both FIGURE 8(a) and FIGURE 8(b). In terms of the average Silhouette coefficient, the value is relatively high when $K = 4$ for both weekdays and holidays. Therefore, comprehensively considering the values of the Silhouette coefficient and SSE, $K = 4$ is the optimal number of the non-checkerboard co-clusters.

Based on the refining operation, the new co-clusters discretize the checkerboard co-clustering patterns into a non-checkerboard co-clustering pattern, which is shown in FIGURE 9. The four classes of co-clusters exhibit different arrival states and are categorized as: ''Very Low'', ''Low'', ''Medium'' and ''High'', which is based on the drop-off volume and the results of K-means. The non-checkerboard co-clusters can reveal more distinctive spatio-temporal patterns, which aid in understanding and detecting differences in the spatio-temporal patterns of human mobility between weekdays and holidays, the results can be drawn as follows.

1) SPATIAL PATTERNS AND THE CORRESPONDING TEMPORAL DYNAMICS

For the non-checkerboard co-clustering patterns, grid cells with the same variation over the entire study area compose a unique spatial pattern. Due to the number of timestamp-clusters in the checkerboard co-clustering, a maximum of four spatial patterns are observed for both weekdays and holidays. According to the composition of each timestamp-cluster, four unique spatial patterns are observed for both weekdays and holidays, which are represented by the green dashed vertical lines in FIGURE 9. The spatial patterns, which were named SPOW-1 to SPOW-4 for weekdays and SPOH-1 to SPOH-4 for holidays, are shown in FIGURE 10; their temporal dynamics are shown as the linear timeline in FIGURE 11.

For the spatial distribution with the ''High'' travel state (red areas in FIGURE 10), it can be observed that the intensive travel state exists in three kinds of spatial

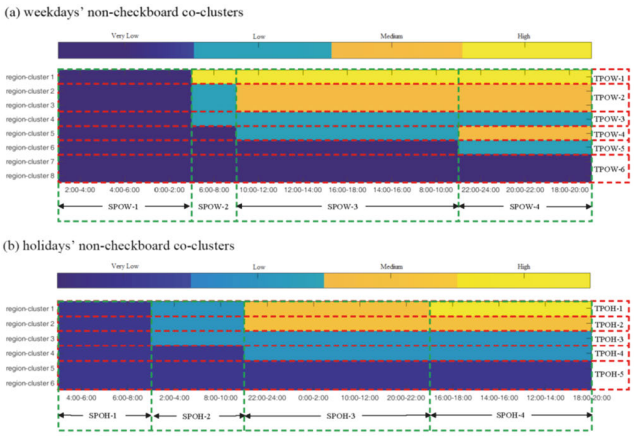


FIGURE 9. Non-checkerboard co-clusters that consist of four arrival states: "Very Low", "Low", "Medium" and "High". The x-axis and y-axis in (a) represent the same checkerboard co-clusters shown in FIGURE 4. The coordinate axis in (b) and the coordinate axis in FIGURE 5 are identical.

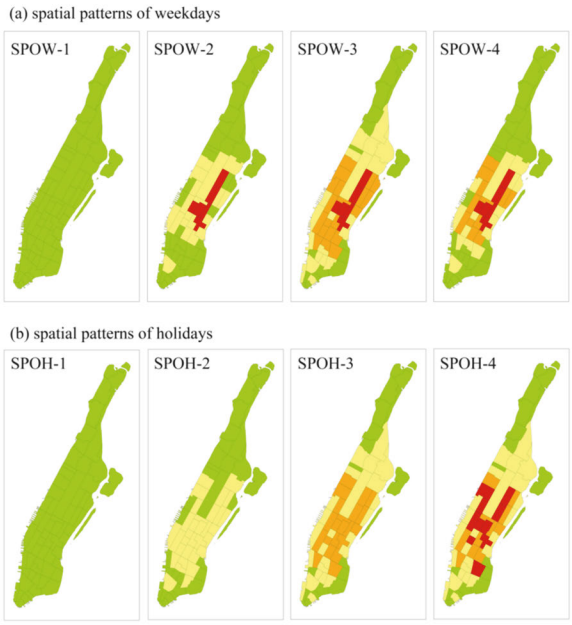


FIGURE 10. Four unique spatial patterns of non-checkerboard co-clusters for both weekdays (a) and holidays (b). The different colors represent the particular travel state: green indicates the "Very Low" state, yellow indicates the "Low" state, orange indicates the "Medium" state and red indicates the "High" state.

patterns on weekdays (SPOW-2, SPOW-3, and SPOW-4 in FIGURE 10(a)) and the corresponding spatial distributions in these spatial patterns are equivalent, which mainly concentrated on Upper East Side and the core regions of Midtown. This finding indicates that travel destinations on weekdays are more aggregated and frequently appear in these specific regions, it is because the diversity of urban functions in these regions, not only has abundant iconic buildings and shopping centers but also has the most luxurious residential district and the largest central business district. However, only one spatial pattern for "High" on holidays (SPOH-4 in FIGURE 10(b)) exists. Compared with the travel destination on weekdays, intensive travel on holidays occurs not only in the Upper East Side and the core regions of Midtown but also in the Upper West Side and some special regions in the Downtown, which reflects the leisure and entertainment functions in these regions. For example, the Upper West Side and the Greenwich Village in the Downtown, which are the holy land of the art and culture in the world, tend to attract more taxi trips on holidays. In particular, there is a unique region in Midtown with a "High" travel state on holidays, which corresponds to the busiest passenger transportation facility in the United States, named Penn Station. The result demonstrates that the rail station plays a more vital role in absorbing activities on holiday than that on weekdays.

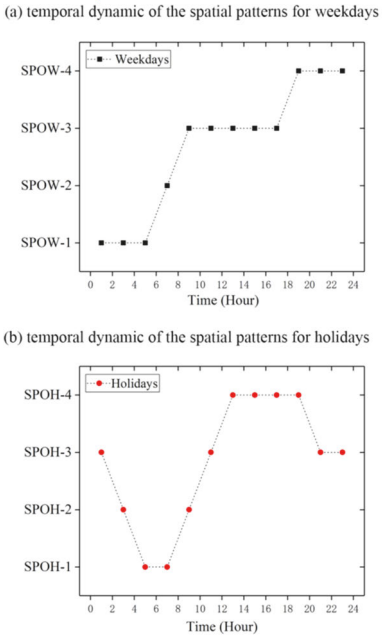


FIGURE 11. The linear timeline shows the temporal dynamics of the spatial patterns. The y-axis shows the spatial patterns in FIGURE 10. (a) weekday timeline; (b) holiday timeline.

In terms of the travel state for "Very Low" (green areas in FIGURE 10), it is observed that a spatial pattern which only "Very Low" travel state for the entire area on both weekdays and holidays (SPOW-1 and SPOH-1), the corresponding time period is [0:00-6:00] for weekdays and [4:00-8:00] for holidays. The temporal distribution of the late-night on holiday fully embodies the title of New York that never sleeps. For other spatial patterns, the spatial distributions with the "Very Low" arrival trips show little difierence between weekdays and holidays in the non-nighttime period, which mainly concentrated on the areas of Upper Manhattan, east of lower Manhattan. The data analysis shows that there is over one order of magnitude difference in the number of trips to these regions compared to other regions, the reason for the above phenomenon may cause by social inequalities, which had been proven in the previous research[57].

Moreover, comparing the travel state distribution in spatial patterns on weekdays, there exhibits greater variance

FIGURE 12. Six unique temporal patterns of non-checkerboard co-clusters for weekdays. The y-axis represents the travel states.

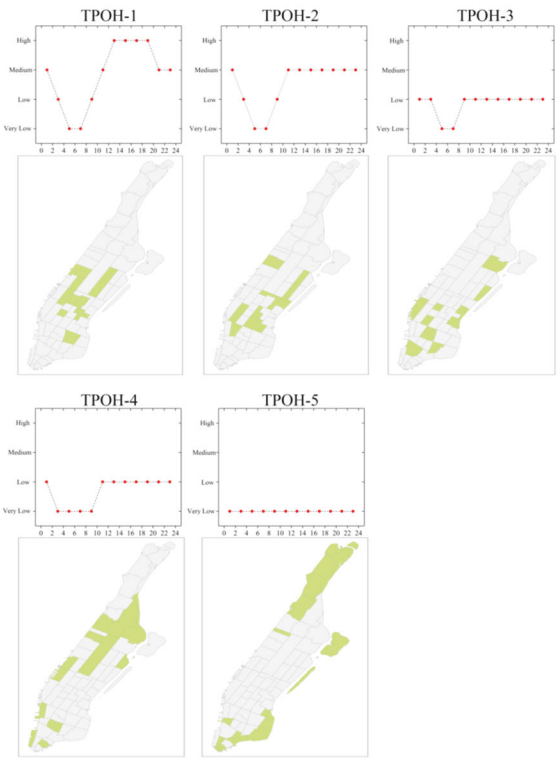


FIGURE 13. Five unique temporal patterns of non-checkerboard co-clusters for holidays.

than the counterpart on holidays, it suggests that the rhythms of people's activities on weekdays vary considerably from place to place while it is more universal on holidays.

The timelines in FIGURE 11 depict the temporal dynamics of the four spatial patterns for weekdays and holidays respectively. On weekdays, the spatial patterns of taxi activities convert by SPOW-1 – SPOW-2 – SPOW-3 – SPOW-4, whereas the spatial patterns on holidays convert by SPOH-3 – SPOH-2 – SPOH-1 – SPOH-2 – SPOH-3 – SPOH-4 – SPOH-3, which represent the more frequent dynamic changes in the spatial patterns. However, the transfer between spatial patterns on holidays is smoother than that on weekdays, the entire region and local regions gradually transfer by ''Very Low'' – ''Low'' – ''Medium'' – ''High'' among different time periods (SPOH-1∼ SPOH-4). In terms of weekdays, some local regions change from the ''Very Low'' state during midnight hours (SPOW-1 in FIGURE 10(a)) to the ''High'' state in the morning hours (red area of SPOW-2 in FIGURE 10(a)), and some regions during the evening peak hours with the ''High'' and ''Medium'' states (red and orange area of SPOW-4 in FIGURE10(a)) convert to the ''Very Low'' state at midnight (SPOW-1 in FIGURE 10(a)). Generally, the results of spatial patterns and the corresponding temporal dynamics indicate a more dynamic yet rhythmic taxi passengers' mobility on weekdays, while a more static yet random structure on holidays.

2) TEMPORAL PATTERNS AND THE CORRESPONDING SPATIAL DISTRIBUTIONS

Similar to spatial pattern partitioning, grid cells with the same variation over the whole day in FIGURE 9 comprise six unique temporal patterns and five unique temporal patterns for weekdays and holidays, respectively. The corresponding temporal patterns are represented by the red dashed transverse lines in FIGURE 9. The temporal patterns are denoted TPOW-1 to TPOW-6 for weekdays in FIGURE 12 and TPOH-1 to TPOH-5 for holidays in FIGURE 13. In FIGURE 12 and FIGURE 13, each timeline indicates one temporal pattern, which varies among the different arrival states over the time period of an entire day, and the highlighted regions in the map represent the geographical extent that corresponds to the temporal patterns. By comparing the timeline between weekdays and holidays, we observe some special temporal patterns.

First, a temporal pattern is observed with only the ''Very Low'' travel state for the entire day on both weekdays and holidays (TPOW-6 in FIGURE 12 and TPOH-5 in FIGURE 13). Although the regional clustering process for weekdays and holidays is independent, the geographical extents are the same for the ''Very Low'' state and the largest among all geographical extents. The quite small volume of taxi activities in these regions may be due to the social inequality analyzed above, such as less economic incentive for taxis to these regions.

TABLE 2. Comparisons with other methods based on Silhouette coefficient, CH, and DBI.

Methods	Weekdays			Holidays		
	Silhouette	CH	DBI	Silhouette	CH	DBI
Co-cluster	0.782	81.849	0.457	0.816	88.054	0.422
K-means	0.638	64.661	0.617	0.820	88.111	0.431
FCM	0.785	75.383	0.501	0.770	80.264	0.579
SOM	0.796	80.924	0.449	0.704	75.573	0.575

Second, TPOH-1 has the most intensive variations in travel state on holidays and converts among four arrival states, which may be attributed to the complex built environment in this geographical extent. Because taxi destinations are related to specific activities [58], regions with different urban functions tend to exhibit different travel patterns with a higher diversity of temporal patterns.

Moreover, we observe that all temporal patterns that change among different arrival states are continuous on holidays, whereas some complex temporal patterns are observed on weekdays. The travel state in the temporal patterns jumped during the evening peak hours and the midnight hours. For example, TPOW-1 converts from ''Very Low'' to ''High'' during the early peak period, largely because of work commuting activities on weekdays, which conforms to the common knowledge of human mobility on weekdays. The jumping change of state indicates that a large spatial variation of taxi activities during the corresponding time period, this finding can be a reference for the taxi company to perform vehicle scheduling.

E. DISCUSSION

1) COMPARE WITH OTHER METHOD

In this section, we compare co-clustering with several traditional clustering algorithms, including the classic k-means [59], ISODATA [60], and SOM [61]. The taxi trip datasets for weekdays and holidays are used to conduct the comparison. We directly cluster spatio-temporal co-occurrence matrices using these three methods by simplifying the spatial and temporal dimensions respectively. Then, we also use the k-means algorithm to regroup the checkerboard-like clusters for all methods. The optimal number of clusters for each method is determined by the Silhouette coefficient and SSE.

Three metrics are calculated to evaluate the quality of cluster results: Silhouette coefficient, Calinski-Harabasz index (CH), and Davies-Bouldin index (DBI). The larger the Silhouette and CH value, the better the clustering effect, while the DBI is the opposite. Table 2 lists the comparison results. Our method, co-cluster, outperforms all the other methods for nearly overall performance, which has maximum Silhouette coefficient and CH value, and the smallest DBI for both weekdays and holidays.

2) LIMITATION

In our research, the mean and standard deviation are used in K-means clustering to refine the spatio-temporal interaction patterns. Indeed, the use of mean, standard deviation to compare different datasets in statistics needs to be based on the premise of conforming to identically distributed (the distribution profiles of our checkerboard-like co-clusters for both weekdays and holidays are basically consistent and present linear distribution approximately). Therefore, it is necessary to test the applicability of the dataset in advance.

V. CONCLUSION

With the development of the mobility big data mining, the spatio-temporal patterns of urban trips can be utilized to comprehensively understand urban traffic and human activities. This article presents a novel analytical framework based on co-clustering to reveal spatio-temporal patterns of urban mobility. It uses a two-step strategy based on the principle of ''divide and group'' which carries out the k-means clustering to generate the refine spatio-temporal co-clusters after the spatio-temporal co-clustering. Compared with other related methods, this article has the following contributions. Firstly, our method is capable of simultaneously clustering the mobility data along both spatial dimensions and temporal dimensions, making it easy to perceive the global similarity in spatio-temporal data. Traditional spatio-temporal clustering methods[6, 7] usually focus on only one dimension (spatial or temporal) or separately analyzes the dimensions. More importantly, the proposed framework provides a concise and efficient way to present the interaction characteristic among spatial and temporal patterns, rather than interactions represented by the combined probability across different patterns [34]. Last but not least, by using the two-step strategy, the method enables the full exploration of the spatiotemporal patterns hidden in the mobility data, thus, the spatial patterns and their temporal dynamics as well as temporal patterns and their spatial dynamics can be discovered.

A case study of taxi data in Manhattan shows that the proposed analytical framework outperforms in identifying the spatio-temporal co-clusters on both weekdays and holidays. Meanwhile, the empirical results uncover some unique spatial and temporal patterns of the urban internal mobility, as well as the interaction characteristics between these patterns and people's activities. The mobility patterns generally indicated distinct spatial activity and temporal dynamic characteristics between weekdays and holidays. Concretely, we find that there is a more dynamic yet rhythmic taxi passengers' mobility on weekdays, while a more static yet random structure on holidays. Moreover, combining the geographical and social background, we observed some particular regional functions implied in the spatio-temporal patterns. For example, rail station plays a more vital role in absorbing activities on holiday than that on weekdays. These findings have important implications for public policies that target specific temporal and geographical units and can be adopted to guide urban infrastructure and transportation planning.

Future improvements to this study could include the following aspects. First, available data could be obtained from the taxi trip datasets in NYC to focus on fixed traffic districts.

Future research should also extend our framework to smaller-scale research areas to reveal the mobility patterns in a fine-scale geographic context. Equally, mining universal travel patterns and abnormal travel characteristics from a more refined time scale is also the direction of future work. In addition, we plan to extend the two-dimensional framework to multiattribute dimensions to detect and explore more complex mobility characteristics. For example, the attribute data of mobility individuals (age and gender), building environment characteristics of the destination, and meteorological factors could be incorporated, which would allow for a deeper exploration of the relationship between urban structure and residents' mobility patterns. We also intend to empirically examine our method using other types of mobility data, such as public bike-sharing data, to explore the universal mechanisms of human mobility in urban areas.

ACKNOWLEDGMENT

The authors thank the Taxi & Limousine Commission of New York City for providing the data presented in this article. The authors would like to thank the editors for the editing assistance and the reviewers for their insightful comments and suggestions.

REFERENCES

[1] G. Atluri, A. Karpatne, and V. Kumar, "Spatio-temporal data mining: A survey of problems and methods," *ACM Comput. Surv.*, vol. 51, p. 41, Sep. 2018.

[2] H. Taubenböck, M. Wiesner, A. Felbier, M. Marconcini, T. Esch, and S. Dech, "New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data," *Appl. Geogr.*, vol. 47, pp. 137–153, Feb. 2014.

[3] X. Wu, R. Zurita-Milla, and M.-J. Kraak, "A novel analysis of spring phenological patterns over europe based on co-clustering: Co-clustering European spring phenology," *J. Geophys. Res., Biogeosci.*, vol. 121, no. 6, pp. 1434–1448, Jun. 2016.

[4] X. Shi, F. Lv, D. Seng, B. Xing, and J. Chen, "Exploring the evolutionary patterns of urban activity areas based on origin-destination data," *IEEE Access*, vol. 7, pp. 20416–20431, 2019.

[5] L. Huang, Y. Yang, H. Gao, X. Zhao, and Z. Du, "Comparing community detection algorithms in transport networks via points of interest," *IEEE Access*, vol. 6, pp. 29729–29738, 2018.

[6] J. Tang, F. Liu, Y. Wang, and H. Wang, "Uncovering urban human mobility from large scale taxi GPS data," *Phys. A, Stat. Mech. Appl.*, vol. 438, pp. 140–153, Nov. 2015.

[7] Y. Chen, Z. Zhang, and T. Liang, "Assessing urban travel patterns: An analysis of traffic analysis zone-based mobility patterns," *Sustainability*, vol. 11, no. 19, p. 5452, Oct. 2019.

[8] G. Andrienko, N. Andrienko, G. Fuchs, and J. Wood, "Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 9, pp. 2120–2136, Sep. 2017.

[9] L. Wu, Y. Zhi, Z. Sui, and Y. Liu, "Intra-urban human mobility and activity transition: Evidence from social media check-in data," *PLoS ONE*, vol. 9, no. 5, May 2014, Art. no. e97010.

[10] Z. Fang, S.-L. Shaw, W. Tu, Q. Li, and Y. Li, "Spatiotemporal analysis of critical transportation links based on time geographic concepts: A case study of critical bridges in Wuhan, China," *J. Transp. Geogr.*, vol. 23, pp. 44–59, Jul. 2012.

[11] I. Boyandin, E. Bertini, P. Bak, and D. Lalanne, "Flowstrates: An approach for visual exploration of temporal origin-destination data," *Comput. Graph. Forum*, vol. 30, no. 3, pp. 971–980, Jun. 2011.

[12] W. Huang and S. Li, "An approach for understanding human activity patterns with the motivations behind," *Int. J. Geograph. Inf. Sci.*, vol. 33, pp. 1–23, Feb. 2018.

[13] J. Tang, W. Bi, F. Liu, and W. Zhang, "Exploring urban travel patterns using density-based clustering with multi-attributes from large-scaled vehicle trajectories," *Phys. A, Stat. Mech. Appl.*, vol. 561, Jan. 2021, Art. no. 125301.

[14] J. Shen and T. Cheng, "A framework for identifying activity groups from individual space-time profiles," *Int. J. Geograph. Inf. Sci.*, vol. 30, no. 9, pp. 1–21, 2016.

[15] X. Weng, Y. Liu, H. Song, S. Yao, and P. Zhang, "Mining urban passengers' travel patterns from incomplete data with use cases," *Comput. Netw.*, vol. 134, pp. 116–126, Apr. 2018.

[16] C. T. Kuo, J. Bailey, and I. Davidson, "A framework for simplifying trip data into networks via coupled matrix factorization," in *Proc. SIAM Int. Conf. Data Mining*, 2015, pp. 739–747.

[17] J. A. Hartigan, "Direct clustering of a data matrix," *J. Amer. Stat. Assoc.*, vol. 67, no. 337, pp. 123–129, Mar. 1972.

[18] X. Wu, R. Zurita-Milla, and M.-J. Kraak, "Co-clustering geo-referenced time series: Exploring spatio-temporal patterns in dutch temperature data," *Int. J. Geograph. Inf. Sci.*, vol. 29, no. 4, pp. 624–642, Apr. 2015.

[19] A. Banerjee, I. Dhillon, J. Ghosh, S. Merugu, and D. S. Modha, "A generalized maximum entropy approach to bregman co-clustering and matrix approximation," *J. Mach. Learn. Res.*, vol. 8, pp. 1919–1986, Aug. 2007.

[20] L. Zhang, C. Chen, J. Bu, Z. Chen, D. Cai, and J. Han, "Locally discriminative coclustering," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 6, pp. 1025–1035, Jun. 2012.

[21] I. Dhillon, S. Mallela, and D. Modha, "Information-theoretic co-clustering," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2003, pp. 89–98.

[22] S. Ullah, H. Daud, S. C. Dass, H. N. Khan, and A. Khalil, "Detecting space-time disease clusters with arbitrary shapes and sizes using a co-clustering approach," *Geospatial Health*, vol. 12, no. 2, pp. 210–216, Nov. 2017.

[23] V. Andreo, E. Izquierdo-Verdiguier, R. Zurita-Milla, R. Rosa, A. Rizzoli, and A. Papa, "Identifying favorable spatio-temporal conditions for west Nile virus outbreaks by co-clustering of MODIS LST indices time series," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 4670–4673.

[24] X. Wu, C. Cheng, R. Zurita-Milla, and C. Song, "An overview of clustering methods for geo-referenced time series: From one-way clustering to co- and tri-clustering," *Int. J. Geograph. Inf. Sci.*, vol. 34, no. 9, pp. 1822–1848, Feb. 2020.

[25] J. Lian, Y. Li, W. Gu, S.-L. Huang, and L. Zhang, "Joint mobility pattern mining with urban region partitions," in *Proc. 15th EAI Int. Conf. Mobile Ubiquitous Syst., Comput., Netw. Services*, Nov. 2018, pp. 362–371.

[26] X. Qian and S. V. Ukkusuri, "Spatial variation of the urban taxi ridership using GPS data," *Appl. Geogr.*, vol. 59, pp. 31–42, May 2015.

[27] D. Zhu, N. Wang, L. Wu, and Y. Liu, "Street as a big geo-data assembly and analysis unit in urban studies: A case study using Beijing taxi data," *Appl. Geogr.*, vol. 86, pp. 152–164, Sep. 2017.

[28] H. Tao, K. Wang, L. Zhuo, and X. Li, "Re-examining urban region and inferring regional function based on spatial–temporal interaction," *Int. J. Digit. Earth*, vol. 12, no. 3, pp. 293–310, Mar. 2019.

[29] X. Liu, L. Gong, Y. Gong, and Y. Liu, "Revealing travel patterns and city structure with taxi trip data," *J. Transp. Geogr.*, vol. 43, pp. 78–90, Feb. 2015.

[30] C. Kang and K. Qin, "Understanding operation behaviors of taxicabs in cities by matrix factorization," *Comput., Environ. Urban Syst.*, vol. 60, pp. 79–88, Nov. 2016.

[31] C. Kamga and M. A. Yazıcı, "Temporal and weather related variation patterns of urban travel time: Considerations and caveats for value of travel time, value of variability, and mode choice studies," *Transp. Res. C, Emerg. Technol.*, vol. 45, pp. 4–16, Aug. 2014.

[32] Y. Liu, C. Kang, S. Gao, Y. Xiao, and Y. Tian, "Understanding intra-urban trip patterns from taxi trajectory data," *J. Geograph. Syst.*, vol. 14, no. 4, pp. 463–483, Oct. 2012.

[33] L. Gong, X. Liu, L. Wu, and Y. Liu, "Inferring trip purposes and uncovering travel patterns from taxi trajectory data," *Cartogr. Geograph. Inf. Sci.*, vol. 43, no. 2, pp. 103–114, Mar. 2016.

[34] L. Sun and K. W. Axhausen, "Understanding urban mobility patterns with a probabilistic tensor factorization framework," *Transp. Res. B, Methodol.*, vol. 91, pp. 511–524, Sep. 2016.

[35] S. Kisilevich, F. Mansmann, M. Nanni, and S. Rinzivillo, "Spatio-temporal clustering," in *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach, Eds. Boston, MA, USA: Springer, 2010, pp. 855–874.

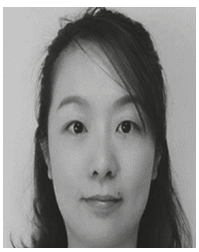
[36] L. Spinsanti and F. Ostermann, "Automated geographic context analysis for volunteered information," *Appl. Geogr.*, vol. 43, pp. 36–44, Sep. 2013.
[37] Q. Liu, W. Liu, J. Tang, M. Deng, and Y. Liu, "Permutation-test-based clustering method for detection of dynamic patterns in spatio-temporal datasets," *Comput., Environ. Urban Syst.*, vol. 75, pp. 204–216, May 2019.
[38] J. Hagenauer and M. Helbich, "Hierarchical self-organizing maps for clustering spatiotemporal data," *Int. J. Geograph. Inf. Sci.*, vol. 27, no. 10, pp. 2026–2042, Oct. 2013.
[39] M. Kulldorff, R. Heffernan, J. Hartman, R. Assuncao, and F. Mostashari, "A space-time permutation scan statistic for disease outbreak detection," *PLoS Med.*, vol. 2, pp. 216–224, Mar. 2005.
[40] S. Shiode and N. Shiode, "Network-based space-time search-window technique for hotspot detection of street-level crime incidents," *Int. J. Geograph. Inf. Sci.*, vol. 27, no. 5, pp. 866–882, May 2013.
[41] Z. Xie and J. Yan, "Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: An integrated approach," *J. Transp. Geogr.*, vol. 31, pp. 64–71, Jul. 2013.
[42] Y. Shi, M. Deng, J. Gong, C. Lu, X. Yang, and H. Liu, "Detection of clusters in traffic networks based on spatio-temporal flow modeling," *Trans. GIS*, vol. 23, no. 2, pp. 312–333, Apr. 2019.
[43] Y. Kluger, "Spectral biclustering of microarray data: Coclustering genes and conditions," *Genome Res.*, vol. 13, no. 4, pp. 703–716, Apr. 2003.
[44] I. S. Dhillon, S. Mallela, and D. S. Modha, "Information-theoretic co-clustering," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2003, pp. 89–98.
[45] H. Cho, I. Dhillon, Y. Guan, and S. Sra, "Minimum sum-squared residue co-clustering of gene expression data," in *Proc. SIAM Int. Conf. Data Mining*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2004, pp. 114–125.
[46] G. Li, "Generalized co-clustering analysis via regularized alternating least squares," *Comput. Statist. Data Anal.*, vol. 150, Oct. 2020, Art. no. 106989.
[47] C. Blöchl, R. A. Amjad, and B. C. Geiger, "Co-clustering via information-theoretic Markov aggregation," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 4, pp. 720–732, Apr. 2019.
[48] N. Andrienko and G. Andrienko, *Exploratory Analysis of Spatial and Temporal Data. A Systematic Approach*. Jan. 2006.
[49] C.-C. Feng, Y.-C. Wang, and C.-Y. Chen, "Combining geo-SOM and hierarchical clustering to explore geospatial data," *Trans. GIS*, vol. 18, no. 1, pp. 125–146, Feb. 2014.
[50] X. Wu, R. Zurita-Milla, E. Izquierdo Verdiguier, and M.-J. Kraak, "Tri-clustering georeferenced time series for analyzing patterns of intra-annual variability in temperature," *Ann. Amer. Assoc. Geograph.*, vol. 108, no. 1, pp. 71–87, Jan. 2018.
[51] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987.
[52] *2018 TLC Factbook*, NYCTLC, New York, NY, USA, 2018.
[53] C. Kadar and I. Pletikosa, "Mining large-scale human mobility data for long-term crime prediction," *EPJ Data Sci.*, vol. 7, no. 1, p. 26, 2018.
[54] Z. Zheng and S. Zhou, "Scaling laws of spatial visitation frequency: Applications for trip frequency prediction," *Comput., Environ. Urban Syst.*, vol. 64, pp. 332–343, Jul. 2017.
[55] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transp. Res. C, Emerg. Technol.*, vol. 36, pp. 1–12, Nov. 2013.
[56] X. Chen, C. Chen, L. Ni, and L. Li, "Spatial visitation prediction of on-demand ride services using the scaling law," *Phys. A, Stat. Mech. Appl.*, vol. 508, pp. 84–94, Oct. 2018.
[57] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, "Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2149–2158, Dec. 2013.
[58] Y. Xu, D. Chen, X. Zhang, W. Tu, Y. Chen, Y. Shen, and C. Ratti, "Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system," *Comput., Environ. Urban Syst.*, vol. 75, pp. 184–203, May 2019.
[59] M. A. White, F. Hoffman, W. W. Hargrove, and R. R. Nemani, "A global framework for monitoring phenological responses to climate change," *Geophys. Res. Lett.*, vol. 32, no. 4, Feb. 2005, Art. no. L04705.
[60] Y. Gu, J. Brown, T. Miura, W. J. Van Leeuwen, and B. Reed, "Phenological classification of the united states: A geographic framework for extending multi-sensor time-series data," *Remote Sens.*, vol. 2, no. 2, pp. 526–544, Feb. 2010.
[61] X. Wu, R. Zurita-Milla, and M.-J. Kraak, "Visual discovery of synchronisation in weather data at multiple temporal resolutions," *Cartograph. J.*, vol. 50, no. 3, pp. 247–256, Aug. 2013.

QIAN LIU received the B.E. degree in electronic information engineering from the University of Science and Technology Beijing, China, in 2011. She is currently pursuing the Ph.D. degree in geographic information science with the School of Information Engineering, China University of Geosciences, Beijing. From 2018 to 2019, she has been co-cultivated with the Department of Physics, Boston University, Boston, MA, USA. Her research interests include spatial data mining and spatial-temporal complex system simulation and modeling.

XINQI ZHENG received the B.S. and M.S. degrees from Henan University, Kaifeng, China, in 1985 and 1988, respectively, and the Ph.D. degree from Information Engineering University, Zhengzhou, China, in 2004. He is currently a Professor with the School of Information Engineering, China University of Geosciences, Beijing. His research interests include spatial analysis and modeling, spatial data mining, spatial optimization, and spatial-temporal complex system simulation and modeling.

H. EUGENE STANLEY received the Ph.D. degree in physics from Harvard University in 1967. He is currently an American Physicist and a University Professor with Boston University, USA. He has made fundamental contributions to complex systems and is one of the founding fathers of econophysics. His current research interests include complexity science and econometrics. He was elected to the U.S. National Academy of Sciences in 2004.

FEI XIAO received the Ph.D. degree from Central South University in 2009. He is currently a Senior Engineer with the Information Center, Ministry of Natural Resources of the People's Republic of China. His current research interest is natural resources big data analysis and applications.



WENCHAO LIU received the Ph.D. degree from the Department of Electronic Engineering, Sheffield University in 2012. She is currently an Engineer with the Information Center, Ministry of Natural Resources of the People's Republic of China. Her current research interest is land–space big data analysis and applications.