Hindawi
Scientific Programming
Volume 2022, Article ID 2955283, 12 pages
https://doi.org/10.1155/2022/2955283

Research Article

An Optimization Model for Landscape Planning and Environmental Design of Smart Cities Based on Big Data Analysis

LiWei Yu,[1] **Lei Zhang**,[2] **and Zhen Gong**[3]

[1]*School of Arts, Qingdao Agricultural University, Qingdao 266109, China*
[2]*Qingdao Institute of Surveying and Mapping, Qingdao 266000, China*
[3]*The Third Branch of Qingdao Institute of Planning and Design, Qingdao 266200, China*

Correspondence should be addressed to Zhen Gong; 200501095@qau.edu.cn

Received 12 March 2022; Revised 2 April 2022; Accepted 5 April 2022; Published 20 April 2022

Academic Editor: Sheng Bin

This paper adopts "digital landscape design logic," analyzes and researches smart city and digital landscape design, and builds a digital city based on design logic, design basis, environment analysis, and results in a presentation based on the dilemma of landscape garden planning and design at this stage and the development trend of the smart garden and digital landscape design. The optimization model of the landscape and environment design is constructed based on design logic, design basis, environment analysis, and result presentation. First, on the Hadoop distributed computing platform based on the MapReduce parallel processing framework, we implement the massive small file processing methods (Hadoop Archives, CombineFileInputFormat, and Sequence Files) to compensate for the inherent defects of Hadoop and experimentally compare the memory consumption and execution efficiency of the three methods to propose a choice. The memory consumption and execution efficiency of the three methods are experimentally compared to propose a selection strategy. Finally, based on the MR-PFP algorithm, we parallelize the frequent itemset in-cab trajectory big data to generate interesting strong association rules. The experimental results show that the MR-PFP algorithm has better speedup ratio performance and higher mining efficiency than the parallel frequent pattern (PFP) growth algorithm. The research and analysis focused on the digital implementation of the standalone environmental analysis, using Rhino software and Grasshopper visual programming language to build parametric logic, establish parametric analysis models, and conduct a comprehensive analysis of the current environment. The study explores the use of digital landscape design methods and technologies in the landscape design process. Using Rhino + Grasshopper parametric and visualization programming software, we built parametric analysis models around elevation, slope, slope direction, water catchment, and viewable area and used mapping and overlay techniques to quantify the urban space. Finally, the purpose of collecting, monitoring, analyzing, simulating, creating, and reproducing landscape information is achieved.

1. Introduction

With the development of computer technology and the Internet, planners have started to use data information to reveal problems in urban space and express the characteristics of urban space and so on. Using data information of urban space as the basis for relevant analysis and research has become the mainstream development direction of urban research now [1]. Data information has become a new language for urban design expression and research. Compared with the traditional graphical urban design language, the data information language is more scientific and objective. However, urban design ultimately needs to be implemented in the three-dimensional space of the city, and the analysis and research results of data information need to be presented in the various spatial diagrams of urban design to guide urban design plans and planning policies more intuitively and effectively [2]. The traditional two-dimensional graphic language can no longer meet the multiplicity and complexity of urban design, and the three-dimensional graphic language, which is used to reveal the mechanism of three-dimensional urban space, has emerged. However, the three-dimensional mapping language is also unable to meet the designers' thinking and description of urban space from more perspectives. Therefore, mapping technology methods and building information models have been applied in urban planning and design.

Big data research is a cross-scientific research method or a knowledge discovery tool, which is closely related to artificial intelligence methods such as data mining, statistical analysis, and search but also differs from the essential connotation of statistics and artificial intelligence [3]. At present, nonrelational data analysis technology represented by Hadoop distributed computing framework and MapReduce parallel programming model has become the mainstream technology of big data analysis under its significant advantages such as being suitable for large-scale parallel processing, unstructured data processing, and simple and easy to use. In the "2015 China Big Data Development Trend Forecast" released at the 2014 China Big Data Technology Conference (BDTC 2014), "Big Data Analytics Combined with Intelligent Computing Becomes a Hot Spot" was selected as the top ten trends in China's Big Data development. The emergence of a huge amount of data storage and computing technology has led humankind into a new era that is different from the past. The new concept of big data is becoming increasingly popular [4]. From Silicon Valley to London, from New York Stock Exchange to Zhongguancun, all industries are paying close attention to the business value, service value, and management value it brings. The emergence of big data processing technology has enabled information, knowledge, and business opportunities that used to be in the hands of a few research institutions to flow out to institutions, organizations, and individuals in the general livelihood sector. Big data technology has been able to develop rapidly precisely because of the essential characteristics of low access threshold, easy access means, and extremely wide coverage, which enable all sectors of society to participate widely.

In recent years, smart gardens and digitalization have now become common in all industries. The rapid growth of digitalization has brought many challenges to urban development. Facing the increasingly complex elements and work of urban gardens, traditional garden design methods depend on the designer's knowledge, experience, and aesthetic ability. This method has been difficult to meet due to the rapid development of cities and the requirements of high standards and scientific urban spatial planning. Modern urban construction urgently needs vitality and innovation, keeping pace with the times, integrating science and technology into disciplinary research, and building a perfect wisdom garden design method [5]. The use of digital technology in urban landscape design realizes the interaction between people and landscape and strengthens the connection between people and landscape, which is a comprehensive, three-dimensional, and interactive process. Based on the characteristics of "digital landscape," the research in this paper takes the urban landscape planning and design process as the mainline and analyzes the applications of landscape data collection, landscape analysis and evaluation, landscape visualization, and landscape construction in the digital landscape. The digital intelligence of each process of urban landscape planning and design improves the science, accuracy, and intelligence of urban landscape planning and design [6]. It provides a theoretical basis for solving the problems of land space utilization, ecological environment, and hydrological system in an urban environment and promotes the construction of the digital city. Further, we explore the prospect of the changes in the forms and ways of public participation under digital technology and the prospects and trends of the use of digital technology in urban landscapes.

2. Related Works

The digital landscape can be broadly divided into digital urban green space, habitat information acquisition and planning, digital landscape system analysis and evaluation, and digital simulation. The International Symposium on Digital Landscape was held on October 14, 2017, at Southeast University on the theme of the digital landscape, where experts and scholars discussed the issues of landscape garden data collection and analysis, sponge city, digital technology, landscape teaching design methods and applications, landscape environment and performance research, landscape environment, and quantitative research [7]. The symposium provided a high-level international academic exchange platform for researchers and practitioners in the urban landscape community to explore the practice of landscape under contemporary digital technology and to focus on the latest research results of the digital landscape at home and abroad today. The promotion of digital technology improves the science and accuracy of urban landscape planning and design and promotes sustainable development. The interactivity, science, and intelligence of the digital landscape complement the new smart city boom [8]. The parallel mining of frequent itemsets is realized by using the MapReduce paradigm, and then interesting strong association rules are generated.

Theoretical research on urban ventilation corridors at home and abroad: research is mainly concentrated in Germany and Japan; Germany is the pioneer in the field of urban wind environment research; after Germany entered the mature stage, Japan began to conduct relevant research and rapid advancement, forming the theoretical basis of ventilation corridors in terms of type, spatial division, mode of action, and guiding recommendations. China's research on urban ventilation corridors is based on German and Japanese research, combined with the actual urban planning system content to put forward the construction of ideas suitable for China [9]. In urban wind environment research and planning spatial scale, it presents three situations, macroscale, microscale research, and research combining macroscale and microscale that scholars have started to carry out in recent years. Big data interaction research at the urban crowd behavior level means using crowd output big data such as hand signaling data, map population density heat to study the behavior and spatial relationships of other people. Based on the above, foreign scholars were the first to carry out this study. Bibri used tens of thousands of GPS devices to obtain activity log data within 2 days and visualized and analyzed the intensity and spatial and temporal transition information of urban daily crowd activities by interacting with urban-scale topographic and planning data [10]. In 2012, Qi et al. analyzed the relationship between employment and residence and commuting in Beijing using LBS transit card data 7 days a week [11]. Adiguzel, Nanjing University, conducted an urban big data study using spatial functional applications based on microblogging platform data [12]. Zhang (2018) applied microblog check-in big data to study the characteristics of regional crowd activities in the Yangtze River Delta [13]. Scholars at home and abroad during the same period obtained the friend relationships between users in social networks in different geographic regions by crawling to call APIs and calculated the strength of connection relationships between urban locations based on the number of friends in different locations. Data information has become a new language for urban design expression and research. Compared with the traditional graphical urban design language, the data information language is more scientific and objective.

Considering the sharing of heterogeneous data from multiple sources and the need to provide data for business systems, the design of smart city system architecture should play an active role in the application and development of smart cities [14]. However, various data-related problems mentioned in the previous section still exist in some current smart city system designs. To solve these problems, this paper designs a data-centric smart city system architecture, which separates the existing data and application support layers and highlights the central position and role of the data layer. This paper also provides a more detailed design of this architecture, considering the need to share multisource heterogeneous data and provide data for business systems, establishing a public data resource platform by establishing a public database, reserving interfaces, or providing data access to various information systems and securing data security by strengthening identity authentication, access rights, data backup, data monitoring, and other measures [15]. The data sharing process is established to ensure that data changes can be tracked and restored, thereby providing higher security assurance for the data sharing process. It is expected that this architecture can provide a reference for the design of subsequent smart city systems.

3. Model Design Based on Big Data Fusion

The recent advances in the technology of storing and mining massive amounts of data for analysis have generated a great deal of interest from governments, research, and business organizations worldwide. There are many web development technologies available, making the web architecture of different websites very different. The most widely used technologies for writing web pages are dynamic java and static HTML or PHP, JSP, XML, and many others, but regardless of the technology used for the architecture, all web pages will eventually transfer data to the user's browser for viewing through the HTTP protocol [16]. Since the behavioral data of city residents required for urban design research is mainly in the form of text values, we mainly use web crawlers to implement it. The crawler system used in the research covered in this paper is based on an open-source project. The crawler implements a distributed operation of the data acquisition task (as shown in Figure 1). The crawler submits the task of obtaining the target information to the server using an object method, and the server obtains the information from the corresponding web page according to the corresponding settings and stores the corresponding source code information of the page. Finally, urban design needs to be implemented in the three-dimensional space of the city. The analysis and research results of data information need to be presented in various spatial diagrams of urban design to guide urban design plans and planning policies more intuitively and effectively. The crawler reads the source code of the corresponding web page, makes all address connections from the source code, and sends the processing results back to the host. Then the program uses the corresponding methods in the object to synthesize and analyze the computational data fed by all the distributed nodes, generates a list of loops for each data acquisition according to a certain strategy, and updates the corresponding indexes, which will be assigned to the distributed crawler in the new loop for the corresponding task.

Find all frequent itemsets from the transaction database whose support is not lower than the minimum support threshold (min_sup) set by the user. That is, $support(i) \geq (min\_sup)$; then, i is a frequent itemset. From the above problem, it can be found that the overall performance of association rule mining depends on problem 1; that is, the key is to find all frequent itemsets. That is, the problem of mining association rules can be reduced to mining frequent itemsets because checking whether they are strong rules is straightforward. Support and confidence are two measures of rule interest, reflecting the usefulness and certainty of the rules found, respectively. A rule $A \Leftrightarrow B$ holds in dataset D with supports and confidence c, where s is the percentage of transactions in D that contain AuB, that is, probability P (A U B); C is the percentage of transactions in D that contain A and also contain B, that is, the conditional probability P(B|A). For example, the support of an item set A is the percentage of transactions in D that contain A. $support(A) = \{T_i | 1 \leq i \leq D\}$ is the relationship between support and confidence

$$\begin{aligned} support(A) &\approx \frac{\text{support\_}count(D)}{\text{count}(A)} + b, \\ \text{confidence}(A \Longleftrightarrow B) &= \frac{\text{support\_count}(A \cap B)}{\text{support}(A)}. \end{aligned} \tag{1}$$

Based on the parallel frequent pattern growth algorithm (PFP), the MR-PFP algorithm introduces a massive small file processing strategy to compensate for the inherent shortcomings of the Hadoop platform in processing massive small file datasets and uses the MapReduce paradigm to achieve the parallel mining of frequent itemsets to generate interesting strong association rules. We evaluate the performance of the MR-PFP algorithm, validate the massive small file processing method and selection strategy, and analyze the experimental results. Specifically, we evaluate the

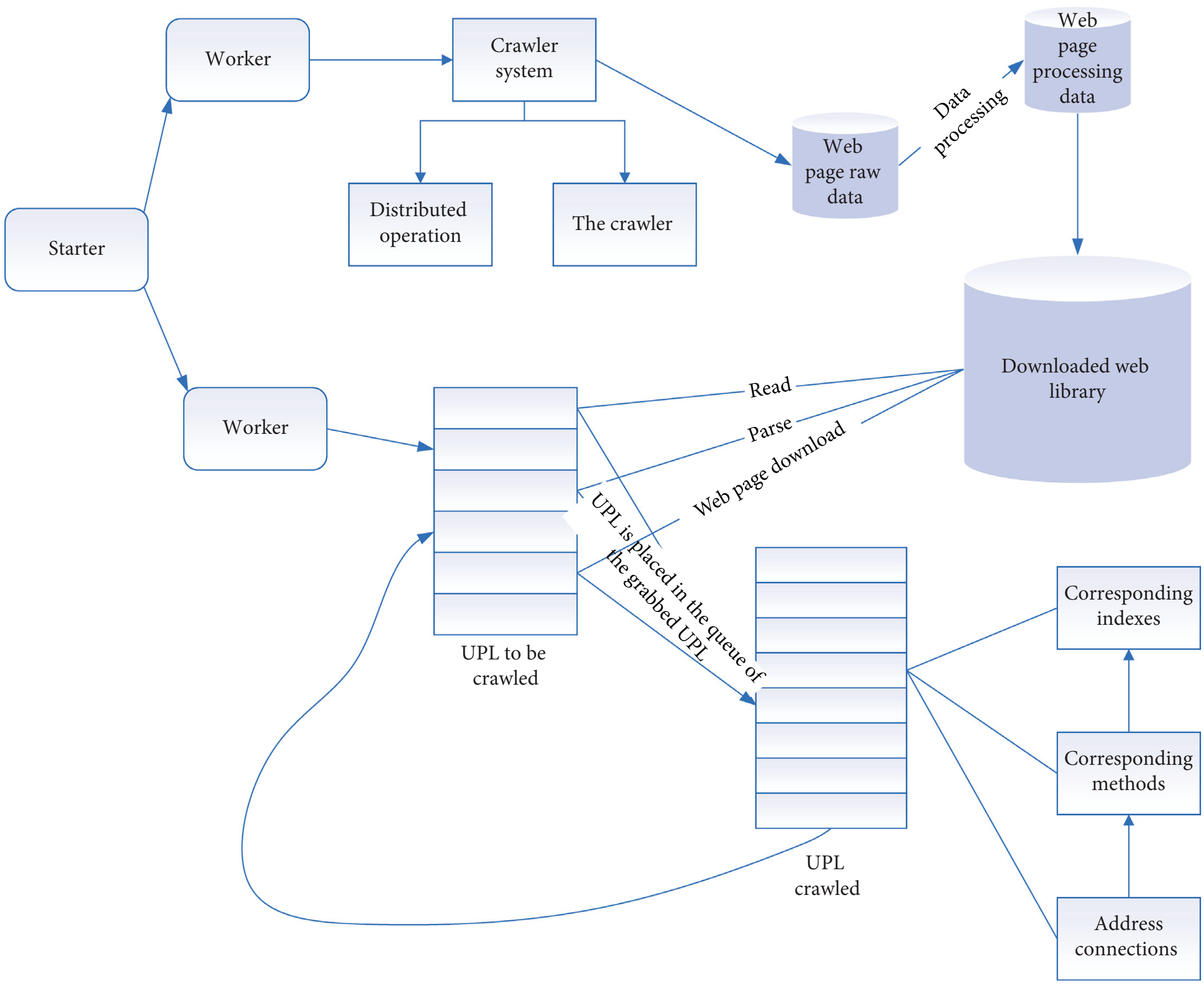

Figure 1: Framework of working principle of web crawler.

mining efficiency and acceleration ratio performance of the MR-PFP algorithm and verify the effectiveness of the massive small file processing method and the rationality of the strategy selection based on the built big data analysis platform with a real massive small file dataset [17]. Based on the built big data analysis platform, the real massive small file data set is used to evaluate the mining efficiency and speedup performance of the MR-PFP algorithm and to verify the effectiveness of the massive small file processing method and the rationality of the strategy selection. In the performance evaluation of the MR-PFP algorithm, the acceleration ratio and processing efficiency are used as evaluation metrics and are compared with the PFP algorithm in the same environment. First, the acceleration ratio performance of the MR-PFP algorithm is evaluated by increasing the number of data nodes in the cluster (i.e., the number of nodes) while keeping the size of three real datasets such as Dataset1, Dataset2, and Dataset3 constant, and the experimental results are shown in Figure 2(a). Secondly, the same dataset is used in the same environment and compared with the PFP algorithm to evaluate the processing efficiency of the MR-PFP algorithm, and its experimental results are shown in Figure 2(b).

Parallel and distributed computing provide powerful techniques for deep mining and efficient analysis of mobile trajectory big data. In this chapter, firstly, the mainstream technologies of parallel and distributed computing are outlined. Secondly, we build a Hadoop distributed computing platform based on the MapReduce parallel processing framework (referred to as "Big Data Analytics Platform"). Finally, the performance of the platform is tested with large-scale real data sets, and the reliability, fault tolerance, scalability, and extensibility of the platform are further verified by expanding applications on the platform. "MapReduce consists of two service modules, JobTracker and TaskTracker, with JobTracker managing the scheduling of jobs (the core of system task allocation) and TaskTracker executing user-defined tasks. MapReduce distributes computational logic to each data node for data computation and value discovery and abstracts the parallel computing process running in a large-scale cluster into Map and Reduce. Its graphic language should also have multidimensional and multilevel expression functions. Therefore, expressions such as mapping technology methods and building information models have begun to be used in urban planning and design. The MapReduce job is partitioned into Map and Reduce

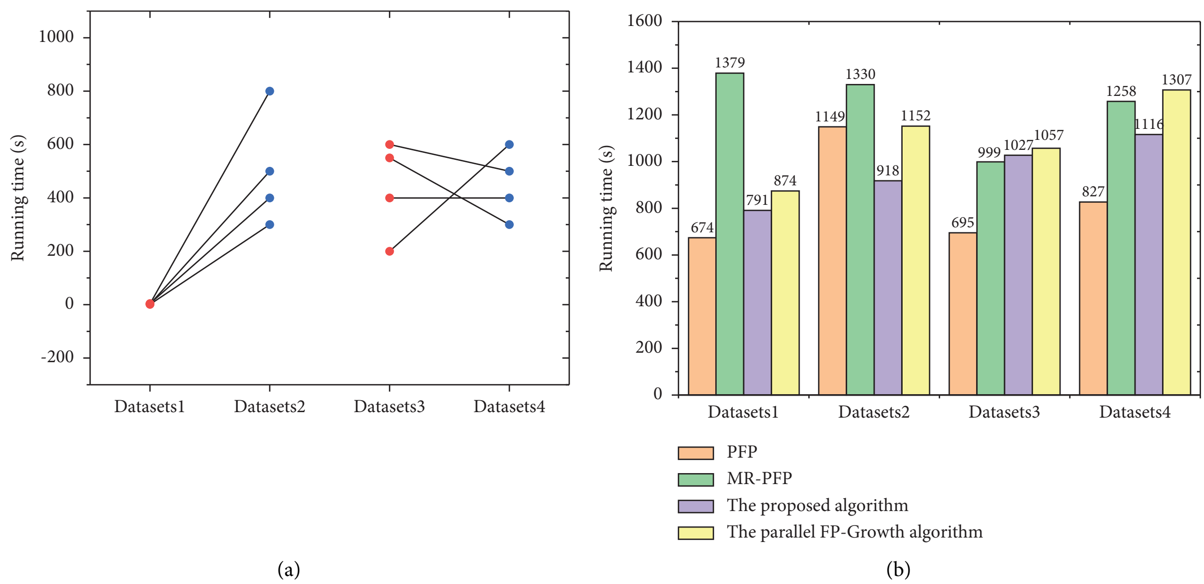

FIGURE 2: Experimental results of MR-PFP algorithm performance evaluation.

tasks and executed in the Map preout phase and Reduce phase, respectively, each with key-value pairs < key, value > as input. The results of the task processing into a new set of intermediate key-value pairs <key, value>, and thus produce the result. That is $support\,(i) \geq (min\_sup)$, then i is a frequent itemset.

$$\mathrm{SSE} = \sum_{K=1} \left\| x_i - uk \right\|, k = \frac{n\sum_{i=1}\left(\hat{X} + x_i^2\right)}{\delta^2}. \tag{2}$$

In the process of building a smart city, there are a wide variety of data sources and many of them, for example, video data from road traffic monitoring in the city and data collected from sensors arranged in the city; how to make these multiple sources of heterogeneous data correlate and interoperate is the focus of data fusion research. Data fusion is an interdisciplinary technology that is characterized by diversity in its research methods regarding data processing. The whole process of data fusion processing and analysis involves many different scientific fields, from acquiring data information to the comprehensive and comprehensive processing of these data. Data fusion is the use of modern computer science to analyze and synthesize the collected or observed information and data with predesigned guidelines to obtain the desired decision or result description. The information samples to be processed for data fusion do not necessarily need to come from the same data source, and scientific research methods from different fields can be utilized in the processing of the data. Modern urban construction is in urgent need of vitality and innovation; keeping pace with the times, integrating science and technology into disciplinary research, building a perfect smart garden design method, and improving the level of informatization, intelligence, refinement, science, and humanization of gardens should be the work of landscape architecture. The principle of data fusion to process data is rough as follows: first, different types of data information are collected through sensors. Secondly, the data feature vectors obtained from the sensors are extracted. Again, the data feature vectors from the previous step are analyzed and judged by statistical and other algorithms through the pattern recognition process; then, the collected data are grouped into target groups, and this grouping process is the data association process; finally, the data are synthesized, and this synthesis process is not just a combination of data, but a complex data fusion process, and the processed data will have a new consistent interpretation or description. The processed data will have a new consistent interpretation or description. The principle of data fusion is shown in Figure 3.

For urban designers, easy-to-use data processing tools can make data visualization problems simple, bypassing complex big data expertise. Especially when urban designers generally do not master computerized methods of processing big data, they need easy-to-use and strong data processing tools to complete the visualization of data, so that the acquired data becomes meaningful and interactive in front of urban design experts and designers. But no matter which technology is used for the architecture, all web pages will eventually transmit data to the user's browser through the HTTP protocol for the user to browse. Based on the use of data visualization tools, urban designers still need to master interdisciplinary skills, not just independent urban designers or urban data analysts, because the complex problem of urban design requires the discovery of spatial knowledge hidden in the data.

4. Model Construction of Intelligent City Landscape Planning and Environmental Design Optimization

The essence of smart gardening is to rely on humanistic wisdom to use technological wisdom to achieve ecological wisdom. Digitization is a form of embodiment of technological wisdom. In the field of architectural design, the

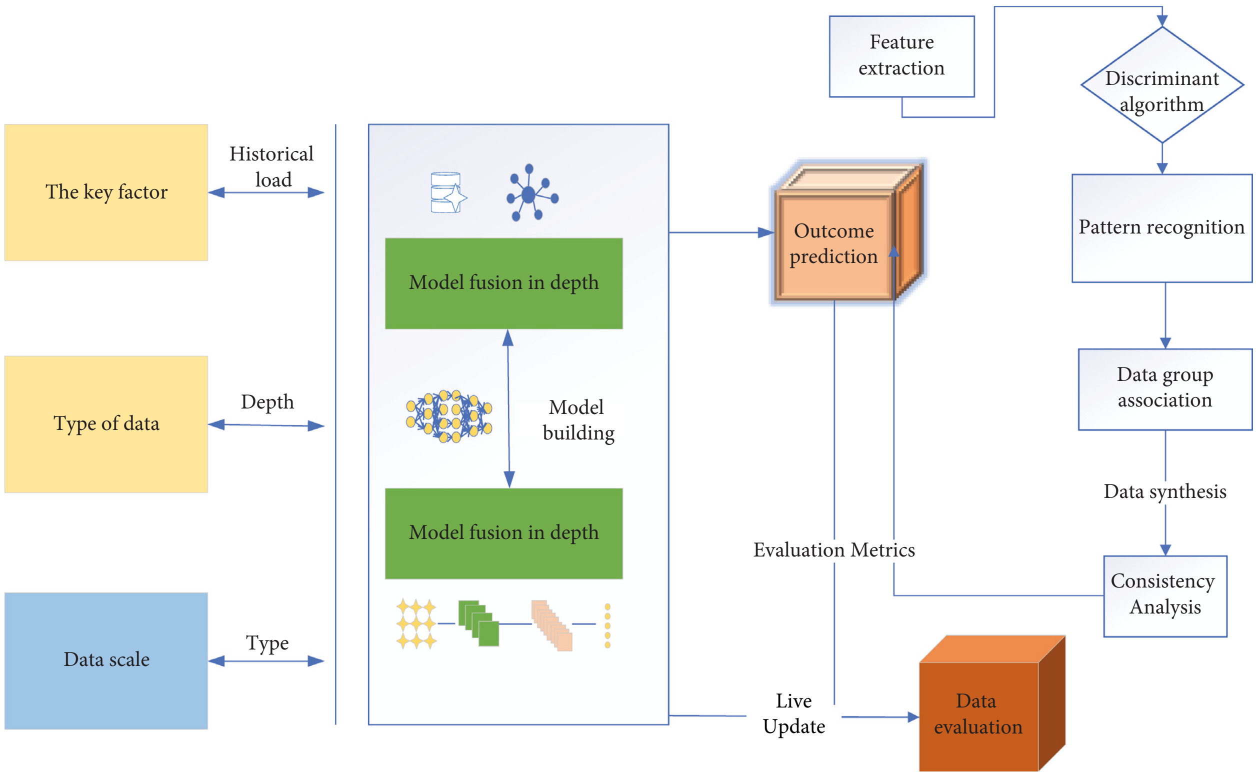

FIGURE 3: Data fusion principal model diagram.

form of buildings is first digitized, using computer programming parametric logic algorithms to form various shaped buildings that are both functional and aesthetic. Landscape design is different from architectural design in that the natural forms of flowers, trees, and flowing water in landscape design elements have a high degree of geometric complexity [18]. Under the existing technical conditions, the amount of data required for the fine depiction of natural forms is very large, and it is difficult to use digital technology as a means of design reproduction only. Computers have the advantage of being able to calculate and quantify, perform several cycles per second, or analyze the relationships between data. Thus, a good research direction for digital landscape design is to use digital for quantitative analysis. The details of design elements are difficult to represent by computer, but their site location, spatial size, scope, landscape elements, or the contours and colors of the space can be digitized. On this basis, topography, light, line of sight, viewable area, and other design factors can be quantitatively described and judged to grasp each element's characteristics and interrelationship. In these aspects of the intervention of digital technology means, then digital in the field of intelligent garden construction can achieve greater value and get a deeper level of development. MapReduce jobs are divided into Map tasks and Reduce tasks and are executed in the Map preexit stage and the Reduce stage, respectively.

The traditional landscape simulation has not adapted to the changing urban spatial environment. The traditional urban landscape simulation is a two-dimensional operation activity, and the urban landscape simulation is mainly shown in the form of a flat surface on a paper drawing. With the development of digitalization and the application of digital technology, designers can use computer composition to express three-dimensional space, turning urban simulation from flat drawings to three-dimensional form, which requires a series of three-dimensional spatial visualization with the help of a variety of digital technologies in the process. For example, virtual reality technology, three-dimensional technology, Unity3D, and other visualization technologies assist landscape designers to visualize the urban landscape [19]. The visualization simulation of urban landscape can be divided into two kinds of landscape process simulation and landscape site simulation, where the urban landscape process simulation is a dynamic simulation different from the static landscape scene. The urban layer system can present 3 types of changes and relationships of urban elements or layers based on the time dimension. First, the urban layer system can elaborate the cotemporal state of multiple urban elements or layers on the same time slice, that is, the presentation of static information of multiple urban elements or layers. Second, the urban layer system can elaborate on the ephemeral evolution of a single urban element or layer with time as a single variable and analyze the relationship between a single urban layer or element and temporal changes. Finally, the urban layer system can also analyze the changing trend of each element or layer in multiple interrelated urban elements or layers with time, the change of internal relationship between elements or layers, and the mechanism of action of the change between elements or layers with time evolution, whose relationship change is shown in Figure 4.

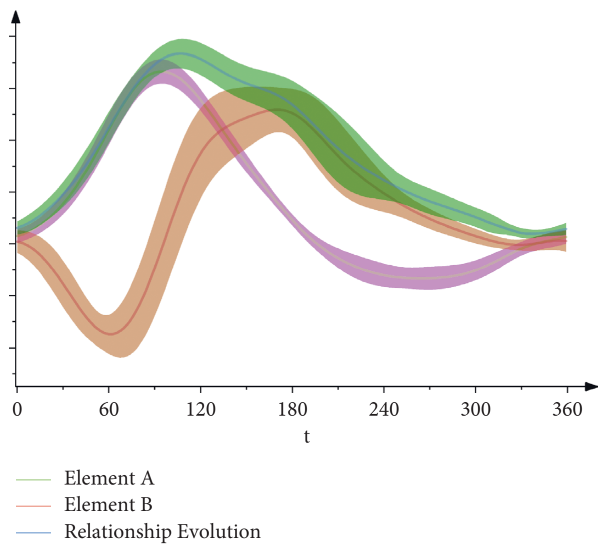

Figure 4: Changes in the relationship between elements or coatings over time evolution.

Urban designers still need to master interdisciplinary knowledge, not just independent urban designers or urban data analysts, because facing the complex problem of urban design requires exploring the hidden spatial knowledge in data. Virtual reality (VR) technology is based on a variety of senses such as human vision, touch, and hearing, and the integrated use of computer technology, multimedia technology, artificial intelligence, and other types of parallel processing. It allows people to be immersed in the virtual world of the computer. Compared with traditional visualization, virtual reality technology has a more realistic experience. The first application of virtual reality in the field of the landscape is the use of 3DMax integrated VRML97 as a means of modeling. Like GIS, VRML97 is a general-purpose software. VRML programs can be directly connected to the web and are supported by many other pieces of software. Unity3D has strong applicability as a game engine in the field of Web3D. Unity3D is popular in many fields due to its simplicity and ease of learning. It plays an increasingly important role in visualizing landscape scenes in urban landscapes. A new virtual reality technology called Augment Reality has been extended based on virtual reality technology. It works by superimposing virtual scenes on the real environment to make it a more realistic space.

The clustering results of the algorithm are reasonable. The clustering analysis results divide the dataset into 7 categories, and the records in each category have a certain similarity. For example, in the category of "front-end technology," the results show that the correlation between HTML5 and Spring Boot is very high, and html5 and other front-end scripting languages are also chosen by most people at the same time. Many people chose both html5 and other front-end scripting languages. In the category of "testing-related technologies," the results also show a high correlation between various testing technologies, which proves the effectiveness of the algorithm. To verify the effectiveness of the improved algorithm, this paper compares the execution time of the Top-K algorithm and the improved algorithm through experimental simulations. The experimental environment is a Windows 10 64-bit operating system with 16 G RAM and CPU Intel i7-9700F 3.00 GHz. The algorithm is run 10 times for the student selection data set, and the running time is recorded. From Figure 5, we can see the trend of the running time of the old and new algorithms as the support increases. Both the old and new algorithms show an upward trend in the overall trend, but the running time of the new algorithm is significantly less than the original algorithm with the same specified support.

To study the effect on the efficiency of the algorithm when the user-set K value increases, we also run the algorithm 10 times, incrementing the K value by 50 each time, gradually increasing the value of K from 50 to 500, and recording the trend of the algorithm's memory consumption. As the support increases, the trend of the running time of the old and new algorithms shows an upward trend in the overall trend. However, the running time of the new algorithm is significantly shorter than the original algorithm when the specified support is the same. As the K value increases, the memory consumption of the algorithm does not change much, which shows the stability of the algorithm. We can also see that the running time of the algorithm increases slightly when the K value is less than about 400, while the running time of the algorithm increases significantly when the K value is greater than 400. This situation may be because the number of user-specified generation rules increases to a certain level, and the algorithm has a sudden increase in the processing time of the data. Therefore, it is recommended not to generate too many rules with this algorithm to avoid the long runtime of the algorithm.

5. Analysis of Big Data Fusion Results

The K value of the K-means algorithm affects the result of clustering. The original K-means algorithm uses the method of randomly selecting K values. This method is extremely unstable and does not give stable and good results. The larger the K value of clustering is, the more detailed the K.1 will be for the analysis of data division, and the sum of squared errors will be gradually reduced. When the K value is smaller than the number of clusters, the SSE (sum of squared errors) decreases because the increase of the K value greatly increases the aggregation of each cluster, and when the K value is larger than the number of clusters, the degree of aggregation will increase if the K value continues to increase, so the decrease of SSE will decrease sharply and then level off gradually. In other words, the K value corresponding to the inflection point of the graph between SSE and K is the optimal number of clusters in the data set.

$$X_k = \sum_{p\in c_i} \left(g(x) + p - m_i\right). \tag{3}$$

SSE is the clustering error of all samples, which represents how well the clustering works. Select the dataset with a certain number of features. Output the corresponding SSE value for each K value. First, calculate the clustering for each K value according to the algorithm, and then calculate each SSE value and save the record. Use the tool to draw the

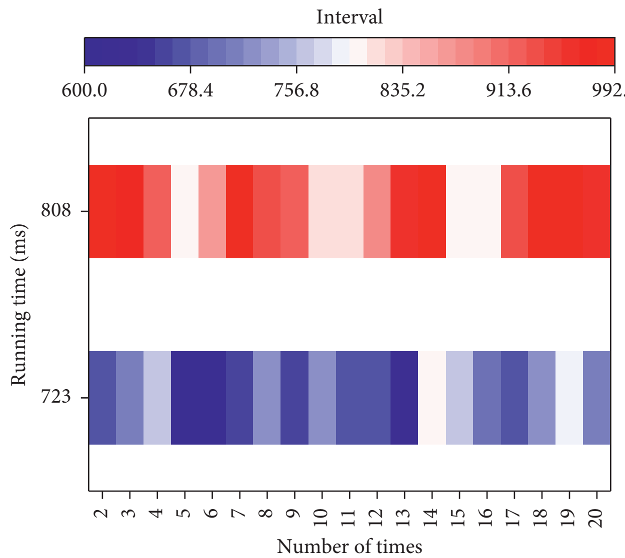

Figure 5: Comparison of the running time of the old and new algorithms.

line graph corresponding to the SSE values corresponding to the K values. Finally, the best K value is determined. The K value determined by the elbow method can be determined as the optimal number of clusters for the current data set. This is more stable and accurate than the original K-means algorithm that randomly selects K values, which effectively avoids the randomness of the K-means algorithm itself and improves the stability of K-means.

$$\sup(r) \leq \min_{\sup(r)}. \tag{4}$$

This may be because the number of user-specified generation rules increases to such an extent that there is a sudden increase in the processing time of the data by the algorithm. Therefore, it is recommended not to generate too many rules when using this algorithm to avoid the running time of the algorithm being too long. The above analysis and discussion are combined to improve and optimize the shortcomings of the K-means algorithm according to the actual needs of this topic, and the stability of the algorithm is improved by the elbow method to optimize the problem that the K-means algorithm cannot determine the K value. In addition, according to the improvement idea of the KNN algorithm, we analyze the difference between the KD-tree structure and the Ball-tree structure, combine the characteristics of the Ball-tree structure and K-means algorithm, and apply the Ball-tree structure to K-means to improve the efficiency of the algorithm running on high-dimensional data. Then, the number of clusters to be specified by the K-means algorithm is determined by the elbow method, and this experiment uses the differential minimum method to determine the number of families K. The evaluation values are recorded for each iteration of X. And the results are plotted as a relation using the elbow method to iterate the algorithm 5–15 times and record the evaluation values for each iteration. The results are plotted as relations. Some of the results are shown in Figure 6.

From the graph and the results after several iterations, we can find that the decreasing trend of the curve starts to slow down when $K = 8$ and $K = 9$, and the decreasing trend becomes flattered and flatter after K is larger than 8 and 9. So we choose $K = 8$, and it is reasonable to cluster the data set into 8 clusters. Then, we run the K-means algorithm, and based on the analysis and discussion above, we decided to use the Ball-tree structure to store the data to improve the efficiency of the K-means algorithm search, reduce the number of calculations, and improve the efficiency of the algorithm.

6. Analysis of the Results of the Application of Smart City Landscape Planning and Environmental Design Model

Since the collected data itself is inaccurate, the act of making judgments with excessive trust in the data, that is, data dictatorship, is extremely dangerous. Whether it is the era of small data or the era of big data, the data collection methods are very different, but the accuracy of the data has the same problem; that is, it may be false, nonobjective, and misleading and not achieve the purpose of quantifying it [20]. Specifically, in urban public space design, whether in the era of small data or big data, designers do not need to conduct on-site research to form an intuitive sense of the design site but can design the area without leaving home by simply consulting various government data annual reports and using Google Earth map data, street views, and photos. Therefore, there are so many cities with no individuality, style, or local cultural and historical characteristics. Many designers have never been to the project site and only rely on CAD maps, site photos sent by others, Google Earth, the last round of planning, and other data to do the design. They hesitate not to feel the local customs and human characteristics on-site, so the projects they make are almost the same, with no local characteristics. This is exactly what urban designers need to strongly avoid when designing spaces and buildings. Meanwhile, security and privacy have been the key issues of big data applications. Using the Top-K algorithm for association rules can reduce the number of data sets and improve the accuracy and quality of association rules mining results.

For parametric analysis, Rhino's modeling capabilities can be used in conjunction with the Grasshopper plug-in's parametric logic algorithm to quantify and simulate topographic and environmental data in the natural environment to guide the design of the project (the base map used for analysis is from "Lan Rhino Art" software education institution). The following analysis of the above-mentioned items introduces the method of parametric analysis in the stand environment analysis. The first thing you need to do is to get and organize the topographic data, there are many ways to get and present the topographic data, and the real purpose is to be able to use such methods to help designers to analyze the site environment, so you can master the methods. This section mainly uses two methods for elevation analysis, and there are three forms of terrain expression: terrain surface viewable,

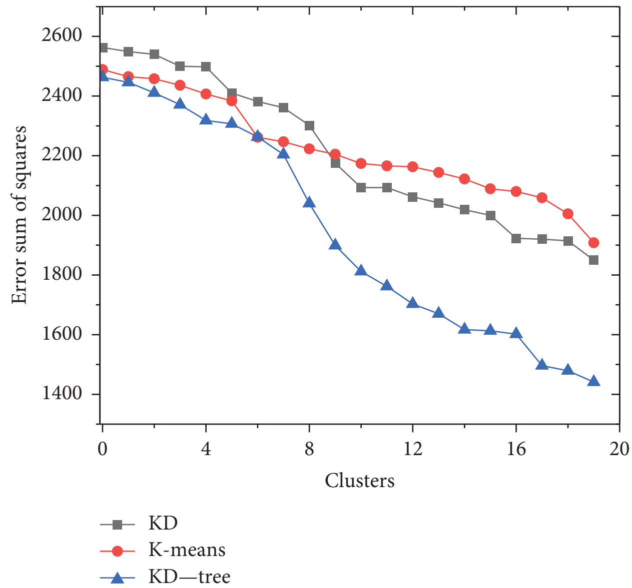

Figure 6: Error sum of squares trend graph.

contour viewable, and elevation point viewable, as shown in Figure 7. In the following section, we will focus on the methods and steps of using DEM data to generate terrain surfaces, contour lines, and elevation points, the application of elevation analysis in scheme design, and the logic construction principles and steps included in the elevation analysis process.

The purpose of the analysis is to classify the users who have the same tendency to choose courses into one category, that is, the same type of course selection among the users in a cluster. The accuracy and quality of association rule mining results can be improved. We normalize the data by discrete normalization, which is essential for the K-means algorithm because the K-means algorithm is calculated by giving Euclidean distance, so if we do not normalize, the larger and smaller numbers will become outliers, and the clustering results will be strongly influenced by these outliers. The choice of the number of nodes in the hidden layer will also affect the training results of the model to a certain extent. In this experiment, the number of nodes in the hidden layer will be determined by an empirical formula. In this experiment, the number of nodes in the hidden layer is determined by the empirical formula. The range of a-values in the formula is 1–10, so the results of different a-values are analyzed, as shown in Figure 8.

To measure the scale growth of the parallel ParKNNO classifier, experiments were conducted by increasing the dataset size (from 1 GB to 5 GB) in the cluster formed by the given nodes, that is, at 2, 4, 6, 8, and 10 nodes, respectively, and the results are shown in Figure 9. From Figure 9, it can be observed that the higher value of scale growth property indicates the longer time consumed when the dataset size is increased in each node. When the dataset size increases from 1 GB to 2 GB, the ParKNNO scale growability value on 2 nodes is close to 2, while the scale growability value on 8 nodes is about 1.5, which is because the communication time of 2 nodes is much shorter than that of 8 nodes. In particular, the communication time in ParKNNO does not change much when the size of the dataset increases exponentially. The results show that the parallel ParKNNO classifier has very good scale growth performance. Based on the above results, it can be found that the parallel ParKNNO classifier based on the MapReduce framework provides an approximately linear speedup ratio, excellent scalability, and scale growth in the Hadoop distributed computing platform. Compared to the serial version corresponding to the centralized standalone environment, the "big data" parallel ParKNNO classifier in a distributed environment obtains higher classification efficiency while producing the same classification results.

In summary, an empirical study based on mobile trajectory big data and the leave-one-out cross-validation (LOO CV) method verify the effectiveness of the TFPC prediction method. The performance evaluation results show that the TFPC method is significantly higher than other methods in terms of prediction accuracy and has a good acceleration ratio, scalability, and scale growth. In other words, the proposed "MapReduce-based traffic prediction method with correlation analysis" achieves better performance and has good robustness in solving the real-life yellow warp planning problem.

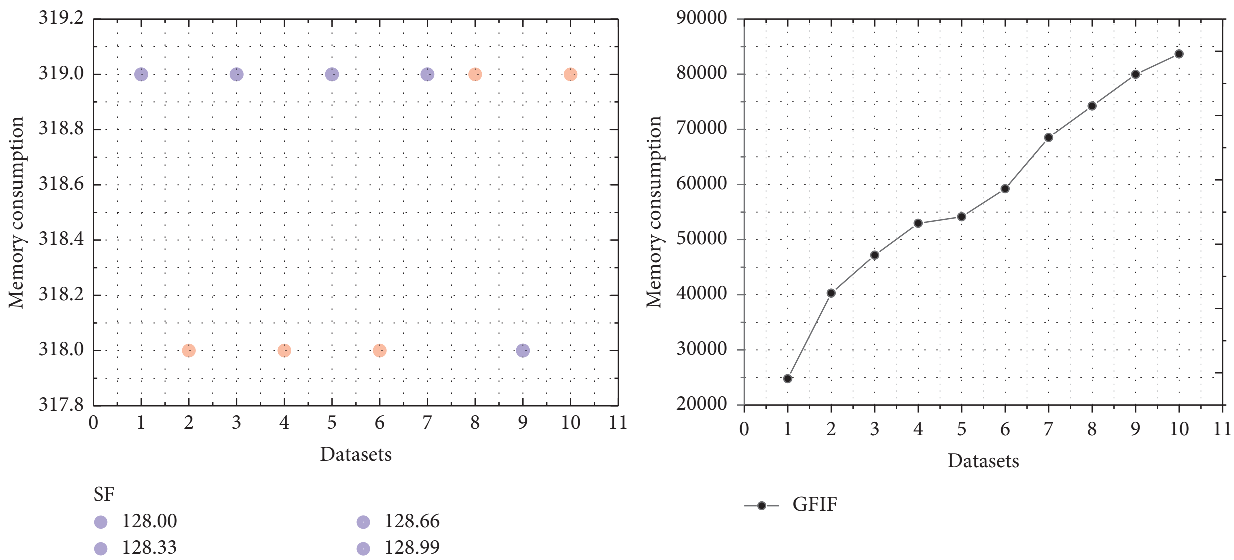

FIGURE 7: Memory consumption of the smart city landscape planning and environmental design model.

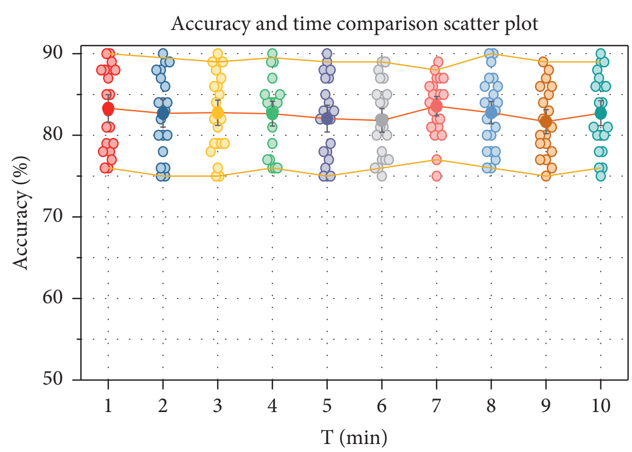

FIGURE 8: Accuracy and time comparison scatter plot.

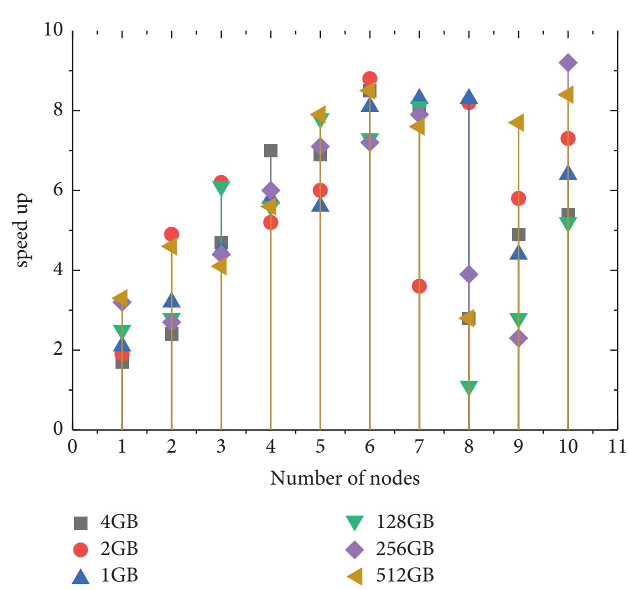

FIGURE 9: Performance results of each index.

7. Conclusion

This paper addresses the problems of real-time, robustness, and accuracy of big data mining, uses big data, data activation, data mining, and other technologies to develop nonrelational data analysis methods in frequent pattern mining and association analysis, cluster analysis, classification, prediction, and so on, deeply mines urban landscape big data, quickly and accurately discovers, refines and analyzes valuable information in trajectory data, and puts forward countermeasures around the government, citizens, and environment highly. We propose countermeasures for the practical problems related to the government, citizens, and the environment and provide decision basis and technical support for building smart cities and then smart cities. The research methods of graphical interpretation, mapping technology, overlay technology, quantitative analysis of urban space, and big data analysis provide the technical basis and methodological support for the construction of a graphical system in urban design. Through Rhino and Grasshopper, we build geographic information spatial models and visualization diagrams, analyze the site environment, master the key information of the site, analyze and extract the explicit and implicit characteristics of the site, discover the problems of the site, utilize the advantages and disadvantages of the site, discover the characteristics of the site, and then design and adjust the plan according to the designer's knowledge and experience.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Project of Special Project of art Education in Shandong Province (Research on Production, Study and Research Teaching Mode of Art Design Major in Agricultural Colleges and Universities under the Background of Rural Revitalization, no. zy20201339) and the Third Branch of Qingdao Institute of Planning and Design(Study on General Survey and Evaluation of Tourism Resources in Jimo District, Qingdao. no. Akk2021-249-YN).

References

[1] E. Carter, P. Adam, D. Tsakis, S. Shaw, R. Watson, and P. Ryan, "Enhancing pedestrian mobility in smart cities using big data," *Journal of Management Analytics*, vol. 7, no. 2, pp. 173–188, 2020.

[2] S. E. Bibri, "A novel model for data-driven smart sustainable cities of the future: the institutional transformations required for balancing and advancing the three goals of sustainability," *Energy Informatics*, vol. 4, no. 1, pp. 1–37, 2021.

[3] K. G. Kim, "Development of an integrated smart water grid model as a portfolio of climate smart cities," *Journal of Smart Cities*, vol. 3, no. 1, pp. 23–34, 2019.

[4] L.-D. Radu, "Disruptive technologies in smart cities: a survey on current trends and challenges," *Smart Cities*, vol. 3, no. 3, pp. 1022–1038, 2020.

[5] T. Deng, K. Zhang, and Z.-J. Shen, "A systematic review of a digital twin city: a new pattern of urban governance toward smart cities," *Journal of Management Science and Engineering*, vol. 6, no. 2, pp. 125–134, 2021.

[6] H. Xu and X. Geng, "People-centric service intelligence for smart cities," *Smart Cities*, vol. 2, no. 2, pp. 135–152, 2019.

[7] M. Coccoli, V. De Francesco, A. Fusco, and P. Maresca, "A cloud-based cognitive computing solution with interoperable applications to counteract illegal dumping in smart cities," *Multimedia Tools and Applications*, vol. 81, no. 1, pp. 95–113, 2022.

[8] Y. J. Dabeedooal, V. Dindoyal, Z. Allam, and D. S. Jones, "Smart tourism as a pillar for sustainable urban development: an alternate smart city strategy from Mauritius," *Smart Cities*, vol. 2, no. 2, pp. 153–162, 2019.

[9] Z. Tong, F. Ye, M. Yan, H. Liu, and S. Basodi, "A survey on algorithms for intelligent computing and smart city applications," *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 155–172, 2021.

[10] S. E. Bibri, "The anatomy of the data-driven smart sustainable city: instrumentation, datafication, computerization and related applications," *Journal of Big Data*, vol. 6, no. 1, pp. 1–43, 2019.

[11] L. Qi, C. Hu, X. Zhang et al., "Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4159–4167, 2020.

[12] F. Adiguzel, M. Cetin, E. Kaya, and M. Simsek, "Defining suitable areas for bioclimatic comfort for landscape planning and landscape management in Hatay, Turkey," *Theoretical and Applied Climatology*, vol. 139, no. 3, pp. 1493–1503, 2020.

[13] L. Zhang, "LA PKU'S research ON China'S traditional eco-wisdom and related application IN contemporary landscape planning and design," *Landscape Architecture Frontiers*, vol. 6, no. 3, pp. 12–28, 2018.

[14] H. S. Shin, "A measure of landscape planning and design application through 3D scan analysis," *Journal of the Korean Institute of Traditional Landscape Architecture*, vol. 36, no. 4, pp. 105–112, 2018.

[15] I. Santé, A. Fernández-Ríos, J. M. Tubío, F. G. Fernandez, E. Farkova, and D. Miranda, "The Landscape Inventory of Galicia (NW Spain): GIS-web and public participation for landscape planning," *Landscape Research*, vol. 44, no. 2, pp. 212–240, 2019.

[16] X. Meng, J. Song, and C. Wan, "The application of marine landscape based on financing model in modern landscape design," *Journal of Coastal Research*, vol. 112, no. SI, pp. 36–39, 2020.

[17] X. Geng, "Rational planning and design of rural road ecological landscape from low carbon perspective," *Ekoloji*, vol. 28, no. 107, pp. 2605–2617, 2019.

[18] K. Raaphorst, “Knowing your audience: the contingency of landscape design interpretations,” *Journal of Urban Design*, vol. 23, no. 5, pp. 654–673, 2018.
[19] P. Shan and W. Sun, “Auxiliary use and detail optimization of computer vr technology in landscape design,” *Arabian Journal of Geosciences*, vol. 14, no. 9, pp. 1–14, 2021.
[20] L. Peng, “Intelligent landscape design and land planning based on neural network and wireless sensor network,” *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2055–2067, 2021.