

## *Retraction*

# **Retracted: Visual Dynamic Simulation Model of Unstructured Data in Social Networks**

## **Security and Communication Networks**

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## **References**

- [1] Z. Xiang, "Visual Dynamic Simulation Model of Unstructured Data in Social Networks," *Security and Communication Networks*, vol. 2022, Article ID 9095330, 10 pages, 2022.

## Research Article

# Visual Dynamic Simulation Model of Unstructured Data in Social Networks

Zhang Xiang 

Natural Sciences Department, General Education College, Jinhua Polytechnic, Jinhua, Zhejiang 321017, China

Correspondence should be addressed to Zhang Xiang; shanon@jhc.edu.cn

Received 29 October 2021; Revised 11 November 2021; Accepted 15 December 2021; Published 13 January 2022

Academic Editor: Jian Su

Copyright © 2022 Zhang Xiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Social networks contain a large amount of unstructured data. To ensure the stability of unstructured big data, this study proposes a method for visual dynamic simulation model of unstructured data in social networks. This study uses the Hadoop platform and data visualization technology to establish a univariate linear regression model according to the time correlation between data, estimates and approximates perceptual data, and collects unstructured data of social networks. Then, the unstructured data collected from the original social network are processed, and an adaptive threshold is designed to filter out the influence of noise. The unstructured data of social network after feature analysis are processed to extract its visual features. Finally, this study carries out the Hadoop cluster design, implements data persistence by HDFS, uses MapReduce to extract data clusters for distributed computing, builds a visual dynamic simulation model of unstructured data in social network, and realizes the display of unstructured data in social network. The experimental results show that this method has a good visualization effect on unstructured data in social networks and can effectively improve the stability and efficiency of unstructured data visualization in social networks.

## 1. Introduction

With the development of computer and mobile terminal technology, Internet-based social network platform is increasingly going deep into people's daily lives, work, and study, and has become the main place and important source of social information diffusion [1–3]. With the continuous rise of online social networking platforms, the unstructured social data are increasing [4]. In social networks, each user is a data source, and its data volume is growing exponentially, and the value is self-evident. In today's information society, whether we can master more information is the key to seizing the market. The platform hopes to obtain useful information from it, but the original data are difficult to understand directly. Data visualization technology can quickly express a wide range of data and visualize information, to reduce the cognitive difficulty of data and help people understand data [5–7]. How to use visualization technology to quickly and effectively find useful data from huge and complex social network data and transform it into

easy-to-understand image information, to serve users and platform managers, has attracted great attention in the industry.

At present, scholars in related fields have studied data visualization and achieved some theoretical results. Kline and Volegov [8] proposed a method to realize 3D data visualization using virtual reality tools. The use of virtual reality technology provides the opportunity to display data, instruments, and experimental settings in three dimensions and allows users to interact with objects, making visualization go beyond two-dimensional projection on a flat screen. Two virtual reality software tools, Unity Gaming Engine and A-Frame, are used for application development to visualize data and high-energy physics targets. This method can effectively use virtual reality tools to realize 3D data visualization. Mrcic et al. [9] proposed the application of social network analysis and data visualization technology in information dissemination analysis. Based on the use of social network analysis, social network data visualization public page links are obtained through Facebook. Through

the development of basic models, data retrieval, data processing, and result analysis and visualization, social network information dissemination is realized. This method has certain validity. However, the above methods still have the problems of poor visualization, stability, and efficiency of unstructured data in social networks.

To solve the above problems, a visual dynamic simulation model of unstructured data in social network is proposed. On the basis of collecting and preprocessing unstructured data of social networks, an adaptive threshold is designed to filter noise, and the visual features of unstructured data of social networks are extracted. On this basis, a univariate linear regression model is established using the Hadoop cluster design and data visualization technology, and HDFS is used for data persistence, MapReduce is used for distributed computing of visual data feature classes, and a visual dynamic model of unstructured data in social networks is constructed to realize the display of unstructured data information in social networks. The social network unstructured data visualization effect of this method is good, stable, and efficient.

## 2. Related Theories and Key Technologies

**2.1. Hadoop Platform.** Hadoop is a distributed system infrastructure, and its origin is an open-source Web crawler Nutch. Hadoop is currently one of the more mature technologies for storing and computing unstructured data and is especially suitable for offline big data processing [10–12]. It is a completely open-source project, running on a large-scale computer cluster, and its core is responsible for storage and calculation.

- (1) HDFS: it is also known as Hadoop Distributed File System. It originated from the Google File System (GFS), which is characterized by distributed processing of data, with the characteristics of large capacity, high reliability, and high availability [13–15]. HDFS supports the storage of very large files, that is, files over hundreds of G, supports streaming data access, and its read-write mode is “write once, read multiple times,” and supports hardware errors. The Hadoop clusters have very low requirements for machine nodes. HDFS data redundancy backup allows hardware errors and is transparent to users. The architecture of HDFS uses a master-slave structure, and the architecture of the HDFS is as shown in Figure 1.

In HDFS, there is one NameNode, one Secondary NameNode, and multiple DataNodes. The default file size in HDFS is 64M. When reading and writing data, the client needs to interact with the NameNode. A Namespace is stored in the NameNode to keep the data storage in the cluster. When storing data, HDFS follows the same data blocks and stores them on different machines and different racks to ensure the effectiveness of data redundancy.

- (2) MapReduce: it is a parallel programming framework for processing unstructured data. It adopts the idea of “divide and rule”, distributes the operation of a large

data set to each node of the cluster, and then collects and integrates the intermediate results of each node to obtain the final results [16–18]. The advantage of MapReduce is that it encapsulates various complex problems in parallel programming. The running process of MapReduce is as shown in Figure 2.

The data transmission between the map task and the reduce task uses a pull model. To be fault-tolerant, the map task stores the intermediate calculation results on the local disk, and the reduce task pulls the corresponding data from each map task through HTTP requests.

**2.2. Data Visualization Technology.** Data visualization is a scientific and technological research on the expression form of data vision. Data visualization includes not only the field of information visualization but also the field of scientific visualization. Data visualization is the visual interpretation of data through modeling, image, animation, and other forms using computer visual image processing [19–21]. Visualization is a display method that transforms abstract things or programs into charts or images. Therefore, data visualization is the way to convert data into charts or images. Data visualization not only expands the traditional display of text, tables, and icons, but also improves the data processing and interpretation, which can better support the decision-making and evaluation of the test process. Therefore, the application of data visualization technology to the research of test data processing and test data visualization, combined with the methods of graphics and images, enables participants to intuitively understand the current situation of testing and strengthen risk prevention, which has important theoretical and practical significance.

Data visualization mostly refers to data visualization and a small part of information visualization. According to different data types and properties, data visualization can be divided into the following types:

- (1) Statistical data visualization: it involves visual display and analysis of statistical data. Most statistical data will be provided in the form of database tables, which are basically used to present data and analyze statistical data.
- (2) Relational data visualization: it is a method used many times in this study, mainly manifested in the relationship between nodes and edges, such as flow chart and network diagram.
- (3) Geospatial data visualization: geospatial data refer to the data describing the specific location of individuals in three-dimensional geospatial space. With the development of the Internet, real-time positioning, map query, route analysis, and other functions are becoming more and more popular, which makes geospatial data more and more important.

Data visualization is not only a collection of algorithms but also a methodological discipline. The visual analysis standard process is suitable for the data visualization analysis process as shown in Figure 3.

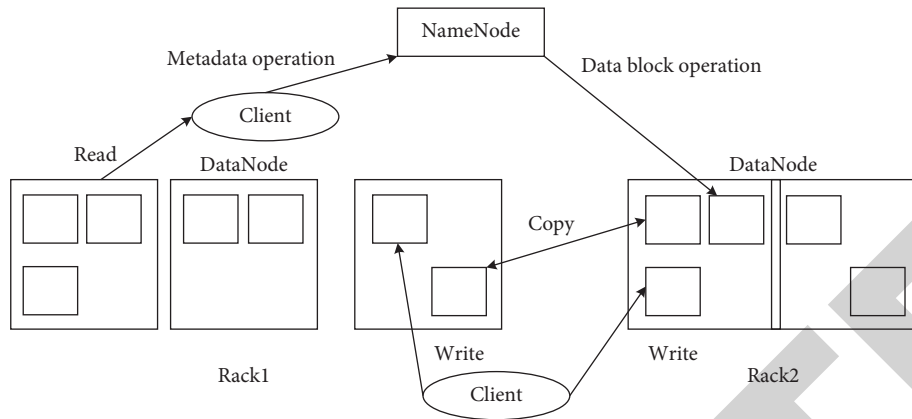


FIGURE 1: HDFS architecture.

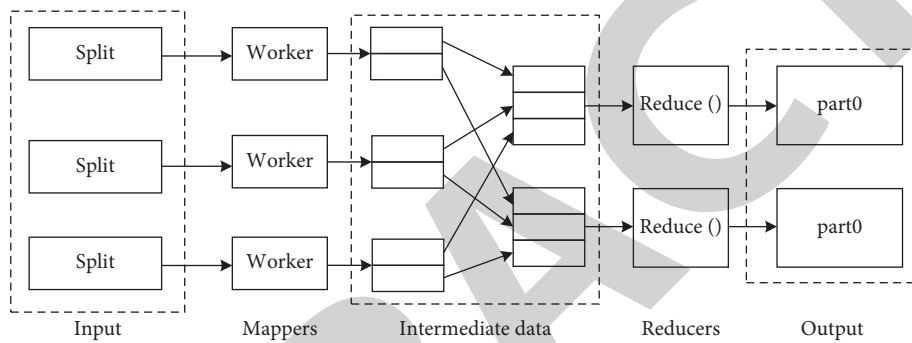


FIGURE 2: MapReduce running program flow.

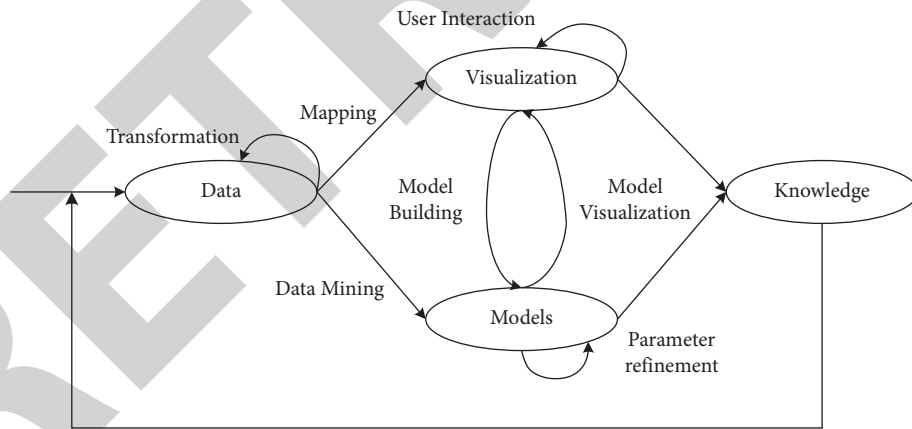


FIGURE 3: Standard flow of visual analysis.

In Figure 3, the input data are the starting point and the output knowledge is the end point. Generally speaking, there are two methods for this transformation process, namely interactive visualization and data mining [22–24]. The intermediate result of this process is the visualization result of data interaction and the data model analyzed from it. It can be seen that the core content of data visualization process includes three aspects: data representation and digital exchange, data visualization representation, and user interaction. How to choose the most appropriate expression form in the complex massive data space will involve the selection of data visualization technology.

Data visualization technology is classified more carefully according to its principle, such as geometry-based visualization technology, icon-based visualization technology, layered technology, 3D technology, and VR technology.

- (1) Geometric projection technology: its basic idea is to use geometric graphics or geometric projection to represent the data in the database. Common technologies include parallel coordinates, scatter plot, and landscape.
- (2) Icon technology: the main logic is to use each part of a simple icon to represent n-dimensional data

attributes. Common techniques include simple strokes and shape coding.

- (3) Layering technology: the main logic is to divide the data space into several subspaces according to the hierarchical structure of the data, which are organized and represented graphically. Hierarchical technology includes tree view, size superposition, and cone tree. The system developed by hierarchical technology mainly includes hyperbolic tree (Xerox), information cube (Sony), and tree map (TreeMap).

### 3. Social Network Unstructured Data Visualization Dynamic Model

This study presents a dynamic visualization model of unstructured data in social networks. According to the time correlation between the data, a univariate linear regression model is established to estimate and approximate the perceptual data. Through the double prediction model, the unstructured data collection model of social network is constructed. According to the collected unstructured data of the original social network, an adaptive threshold is designed to filter the impact of noise. Then, feature analysis is carried out on the processed social network unstructured data, and the visual features of social network unstructured data are extracted. The Hadoop cluster design is introduced, HDFS is used for data persistence, MapReduce is used for data feature extraction and clustering for distributed computing, and multilevel parallel operation of visual feature extraction of unstructured data in social networks is realized.

**3.1. Unstructured Data Collection.** The topological structure of social networks is characterized by chain and long distance, and the distribution of nodes in social networks is sparse [25–27]. According to the connectivity in physical phenomena and time series, there is a time correlation between the data collected in the same node in the same time period. According to the time correlation between the data, a univariate linear regression model is established to estimate and approximate the perceptual data. Through the dual prediction model, that is, the cluster head and nodes in the cluster use the same model to predict the data, reduce the amount of data transmission, build the unstructured data collection model of social network, and complete the collection of unstructured data in social network.

The simplest univariate linear regression model  $\alpha' = a + bt$  is used to avoid the problem of high algorithm complexity caused by the limited storage space and node computing power of social networks and reduce the time used for data collection. Among them,  $t$  represents the time used to collect unstructured social network data,  $\alpha'$  represents the corresponding predicted value of  $t$ , and  $a$  represents the true value of  $t$ . The sampling of nodes in the social network is completed according to the time sequence, and  $n$  data sets are obtained, which are denoted as  $TS$ , and the expression is as follows:

$$TS = \{(t_1, \alpha_1), (t_2, \alpha_2), \dots, (t_n, \alpha_n)\}, \quad (1)$$

where  $TS$  can be used as a function of the dependent variable as the sampling value  $\alpha'$  and the independent variable as the sampling time  $t$ , and the univariate linear regression model can be obtained by the least-squares method [28–30]. To reduce the sum of squared errors between the fitted curve and the sampled data, let

$$D = \sum_{i=1}^n [\alpha_i - (a + bt_i)]^2. \quad (2)$$

The second-order partial derivative of the parameters  $a$  and  $b$  is solved by the function  $D$ , so that the error between the real value and the predicted data is reduced:

$$\begin{aligned} a &= \bar{\alpha} - b\bar{t}, \\ b &= \frac{\sum_{i=1}^n t_i \alpha_i - n\bar{t}\bar{\alpha}}{\sum_{i=1}^n t_i^2 - n\bar{t}^2}. \end{aligned} \quad (3)$$

The nodes in the social network calculated the parameters  $a$  and  $b$  through formula (3) and transmitted them as the parameters of the data acquisition model to the corresponding cluster head nodes in the social network, which reduces the energy consumption and information redundancy in the social network. The cluster head node establishes the data collection model  $\alpha'$  of the social network node according to the parameters  $a$  and  $b$  to complete the collection of big data in the social network:

$$\alpha' = a + bt. \quad (4)$$

The unstructured data collection process of social network is shown in Figure 4.

In Figure 4, the first part is to collect data in the social network through time series. The sensor will not transmit network data before data collection and model establishment. The second part uses the least-squares method to construct the regression model, and the later changes in the data can be predicted by the univariate linear model. The third part is to identify whether the collected data exist in the error range. If the collected data are in the error range, the sampling frequency of nodes in the social network should be adjusted in real time. If the collected data are not within the error range, it is discarded and returned to the first part.

**3.2. Unstructured Data Processing.** According to the collected unstructured data of the original social network, an adaptive threshold is designed to filter the impact of noise. Assuming that the average value of the original signal amplitude collected by a sensor is  $r(n)$  and the superimposed noise is  $u(n)$ , the original signal amplitude can be expressed as follows:

$$|x(n)| = |r(n) + u(n)|, \quad n \in \Omega(r), \quad (5)$$

where  $\Omega(r)$  is the target data distribution range. Assuming that the average value of the noise term  $u(n)$  is 0 and the variance is  $\sigma^2$ , then  $u(n)$  satisfies the  $N(\sqrt{\pi/2}\sigma, (4-\pi)\sigma^2/2)$  distribution. As long as the original signal contains the target signal, a certain fluctuation will occur when calculating the

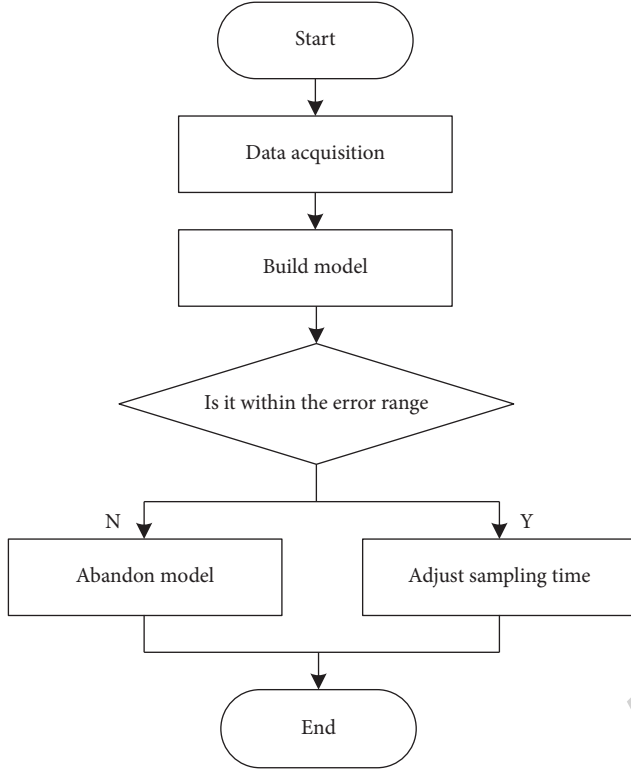


FIGURE 4: Flow chart of social network unstructured data collection.

amplitude difference, and this fluctuation is not related to the amplitude. Therefore, the threshold of the difference signal can accurately filter out the noise doped by the original signal [31–33].

Assuming that the threshold is  $t$ , when determining whether there is a target signal, the error probability is expressed as follows:

$$p = \int_t^{\infty} \frac{\omega}{\sigma^2} \exp\left\{-\frac{\omega^2}{2\sigma^2}\right\} d\omega + \int_0^t \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\omega-r)^2}{2\sigma^2}\right\} d\omega. \quad (6)$$

According to the continuity of the probability density, the partial derivative of  $t$  can be obtained from  $p$ , so that

$$\frac{\partial p}{\partial t} = \frac{t}{\sigma^2} \exp\left\{-\frac{t^2}{2\sigma^2}\right\} + \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(t-r)^2}{2\sigma^2}\right\}. \quad (7)$$

In addition, according to the unimodal characteristic, the partial derivative calculation must have a minimum value, and let  $\partial p/\partial t = 0$ ,  $t = a\sigma$ ,  $r = b\sigma$ ; thus  $\sqrt{2\pi}a = e^{(ab-b^2/2)}$  is solved, and the selection of the threshold  $t$  is transformed into a discussion of the parameter  $a$ . Because the threshold setting should exceed the noise amplitude and not exceed the target signal amplitude, there is  $\sqrt{2\pi} < a < b$ . At the same time, to accurately detect the smallest possible signal, it should be ensured that the amplitude of the target signal does not exceed  $b_0$  times of  $\sigma$ , that is,  $e^{(ab-b^2/2)} \leq e^{b_0^2/2}$ , so the relationship between  $a$  and  $b$  is calculated by the Taylor expansion to obtain:

$$\sqrt{2\pi}\sigma = e^{(ab-b^2/2)} = \sum_{s=0}^{s=b_0} \frac{(ab-b^2/2)^s}{S!}. \quad (8)$$

At this time, only the  $b_0$  value needs to be adjusted according to environmental changes, and the corresponding threshold value can be calculated.

**3.3. Visual Feature Extraction of Unstructured Data.** To complete the visualization of social network unstructured data, we need to do feature analysis on the processed social network unstructured data; that is, we need to do feature clustering extraction of various sensor data as the input of the social network unstructured data visualization dynamic model and convert a large number of social network unstructured data into visual dynamic results by comparing with the eigenvalues of the expert database. When taking feature extraction for a certain type of sensor data, suppose its clustering center is represented as  $C(x_i, A_j(L))$ , where  $x_i$  represents the state  $i$  and  $A_j(L)$  represents the neighborhood range, then the clustering operation can be described as follows in the process of analyzing spectral features:

$$x_m(t) = \sum_{j=1}^I S_j(t)e^{j\phi} + n_m(t), \quad (9)$$

where  $S_j(t)e^{j\phi}$  represents the distribution index of the sensing parameter, and the corresponding sensing parameter will be further characterized according to its range. Assume that the scattered range of multidimensional features is  $w_{is}$  and the source of unstructured data sensing parameters of social network is  $n$ . Therefore, the spectrum  $r_i$  of the visual feature distribution of a certain sensor parameter can be calculated as follows:

$$r_i = \frac{1/n \sum_{i=1}^{n-1} [x_i(s-l) - w_{is}]^3}{(1/n \sum_{i=1}^{n-1} [x_i(s-l) - w_{is}]^2)^{3/2}}. \quad (10)$$

The kurtosis of unstructured data feature distribution of social networks is expressed as follows:

$$k_i = \frac{1/n \sum_{i=1}^{n-1} [x_i(s-l) - w_{is}]^4}{(1/n \sum_{i=1}^{n-1} [x_i(s-l) - w_{is}]^2)^2} - 3. \quad (11)$$

Because the visual features have high-dimensional characteristics, it is necessary to transform the data. Here, the wavelet algorithm [34–36] is introduced to reduce the dimension, and the correlation function between the feature parameters and the distribution map is decomposed, which is described as follows:

$$G_n(r) = \frac{2}{n_2 - n} \sum_{i=1}^n \sum_{j=i+1}^n f(r_i - |x_j|). \quad (12)$$

According to the correlation function, the step size of sensing parameter feature extraction can be obtained. Therefore, the visual feature extraction function is designed as follows:

$$F = \text{amin}_{z=0}^n G_n(r)x_{z\text{step}} + b. \quad (13)$$

Here,  $a$  and  $b$  are feature factors, and effective data for visualization in the monitoring data can be calculated through the feature extraction function. It is assumed that the extracted feature vector is expressed as  $[c_1, c_2, \dots, c_n]$ , and any feature element can be described as  $c_i = F(v = i|u_i)$ . Since the extracted function does not have nonlinearity, reconstruction is adopted here to integrate the statistical features of monitoring data by compressing the features. The process is expressed as follows:

$$p(t) = \sum_i u_i(t)e^{j\phi}c(t - iT_S), \quad (14)$$

where  $u_i(t)$  represents the amplitude of data,  $c(t - iT_S)$  represents the impact response, and  $T_S$  represents the time interval and represents the time density of data. The larger its value is, the more detailed the description of monitoring data is. Its calculation method is shown as follows:

$$T_S = \sqrt{\int_0^T (\tau - \bar{t}_n)^2 [s(t)]^2 dt}, \quad (15)$$

where  $s(t)$  represents the time-domain shape of the data, and  $\bar{t}_n$  represents the average processing time, which can represent the distribution density of data features. Its calculation method is as follows:

$$\bar{t}_n = \int_0^T \tau \frac{[s(t)]^2}{\int_0^T [s(t)]^2 dt} dt. \quad (16)$$

According to the time series, the data features are visually reconstructed and described as follows:

$$M(t+1) = \frac{\sum_{i=1}^n c_i(t)(u_i - p(t+1)(u_i - p(t+1)))^T}{\sum_{i=1}^n X}, \quad (17)$$

where  $X$  represents the feature vector. Using the reconstruction model, nonlinear real-time extraction of social network unstructured data can be performed, which is beneficial to the realization of dynamic visualization.

**3.4. Unstructured Data Visualization Feature Processing Model.** To deal with a large amount of unstructured data generated by social networks and provide dynamic and efficient data processing parameters for visualization, the Hadoop cluster design is introduced, HDFS is used for data persistence, MapReduce is used for data feature extraction and clustering for distributed computing, and cluster design and distributed computing realize the multilevel parallel operation of visual feature extraction, so the accuracy and real-time performance will be improved. The Hadoop visual feature processing model is as shown in Figure 5.

According to the model description, after filtering out the noise in the original data through the acquisition and fusion process, the data are sent to the HDFS as the input of Hadoop. At this time, part of the data and intermediate parameters will be persisted, and part of the data will be used as MapReduce input. MapReduce first adopts singular value judgment for the

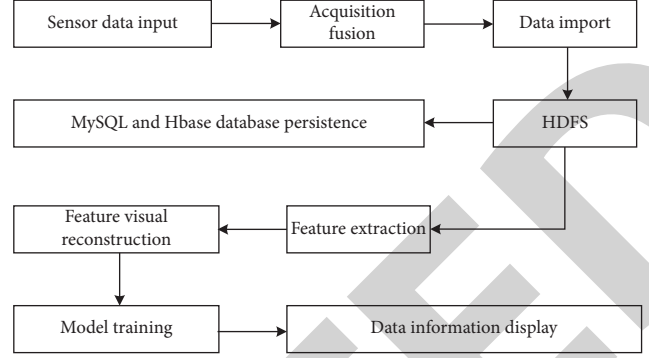


FIGURE 5: Hadoop visual feature processing model.

input and then adopts clustering operation for the data to obtain the visual feature distribution. According to the time-domain and frequency-domain transformation, the wavelet algorithm is used to decompose the correlation function between the feature parameters and the distribution map, to achieve the effect of dimensionality reduction and simplification. For the linear model, the time-series reconstruction visual model is used to extract data features. Finally, a large number of input data are transformed into data information display results after redundancy removal and feature extraction and classification, to realize the visualization of unstructured data in social networks.

## 4. Experimental Analysis

### 4.1. Experimental Environment and Parameter Setting

**4.1.1. Experimental Environment.** To verify the effectiveness of the social network unstructured data visualization dynamic simulation model, in the MATLAB platform, the Windows 7.0 operating system was used to verify the overall performance of the social network unstructured data visualization dynamic simulation model. A Hadoop distributed cluster is built based on the Linux system to realize dynamic visual processing of unstructured data in social networks. 1 master server and 5 slaver servers in the cluster are configured. `hadoop-env.xml`, `core-site.xml`, `hdfs-site.xml`, and `mapred-site.xml` files of each server are independently configured, the corresponding IP address of each server is set to 192.168.1.100–105, and SSH key is used to set up secret-free access for all nodes. The collection and fusion algorithm and the visualization feature processing algorithm are implemented in *Python* and deployed in the Hadoop cluster, and the persistence tool is used in the HBase database.

**4.1.2. Parameter Setting.** To verify the proposed visual dynamic simulation model of unstructured data of social network, Simulink tool is used to read unstructured data information, extract information feature components, and set the node coverage of social network to  $240 \times 240$ , and the node scale is set to 300 cm, the optimal control points are 10, the information communication coverage radius of the output node is  $r=1.5$ , and other simulation parameter settings are shown in Table 1.

TABLE 1: Simulation parameter setting.

Parameter type	Value
Simulation area	240×240
Number of node topologies	20
Routing and forwarding protocol	IEE 802.15.4
Data characteristic distribution variable	1200
Address resolution	12.4
Location of nodes	(0, 0)

4.2. *Visualization Effect Analysis.* To verify the social network unstructured data visualization effect of the proposed method, the data are loaded, the proposed method combined with attribute clustering algorithm is selected, the corresponding layout parameters for visual layout are set, and the social network unstructured data information display results of the proposed method are obtained as shown in Figure 6.

According to Figure 6, the proposed method can effectively realize the visualization of unstructured data of social network. The visualization effect of unstructured data of social networks is good. It can display the unstructured data information of social network, has strong data carrying capacity, and can show the network structure and information distribution.

4.3. *Visual Stability Analysis.* To verify the stability of the proposed method for visualization of unstructured data in social networks, the stability coefficient is used as the evaluation index.  $\kappa$  is set as the stability coefficient of unstructured data visualization of social networks. When the stability coefficient takes a value in the interval  $\kappa = [2, 6]$ , the stability of unstructured data visualization of social networks is relatively high. The method of reference [8] and the method of reference [9] are, respectively, used to compare them with the proposed methods, and the visual stability comparison results of unstructured data of social networks with different methods are obtained as shown in Figure 7.

It can be seen from Figure 7 that when the number of iterations is 500, the stability coefficient of the social network unstructured data visualization method in reference [8] takes a value in the interval  $[1.5, 6.3]$ , the stability coefficient of the social network unstructured data visualization method in reference [9] takes a value in the interval  $[1.8, 6.7]$ , and the stability coefficient of the social network unstructured data visualization of the proposed method takes a value in the interval  $[2.1, 5.5]$ . It can be seen that, compared with the method of reference [8] and the method of reference [9], the stability coefficient of unstructured social network data visualization of the proposed method takes a value in the interval  $\kappa = [2, 6]$ , which can effectively improve the visualization of social network unstructured data stability.

4.4. *Efficiency Analysis.* On this basis, the visualization efficiency of social network unstructured data of the proposed method is further verified. Taking the visualization time of social network unstructured data as the evaluation index, the shorter the visualization time of social network unstructured

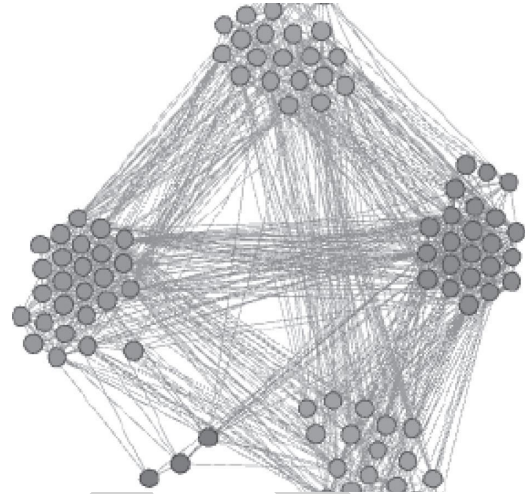


FIGURE 6: Display results of social network unstructured data information of the proposed method.

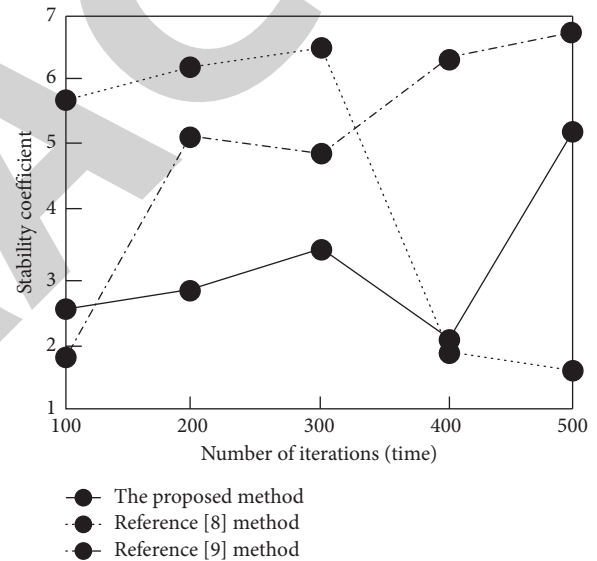


FIGURE 7: Comparison results of visual stability of unstructured data in social networks with different methods.

data, the higher the visualization efficiency of social network unstructured data. By comparing the method of reference [8], the method of reference [9], and the proposed methods, the visualization time comparison results of unstructured data of social networks with different methods are obtained, as shown in Figure 8.

As can be seen from Figure 8, with the increase in unstructured data, the visualization time of unstructured data of social networks with different methods increases. When the amount of unstructured data is 2500 GB, the social network unstructured data visualization time of the method in reference [8] is 56s, and the social network unstructured data visualization time of the method of reference [9] is 42s, while the social network unstructured data visualization time of the proposed method is only 26s. Therefore, compared with the method of reference [8] and the method of



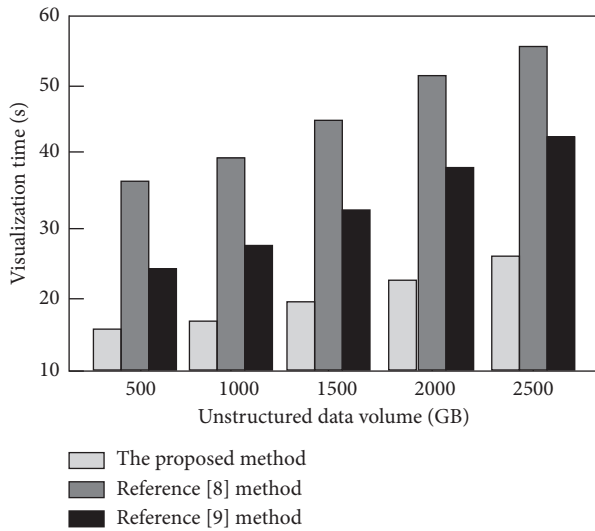


FIGURE 8: Comparison results of different methods of social network unstructured data visualization time.

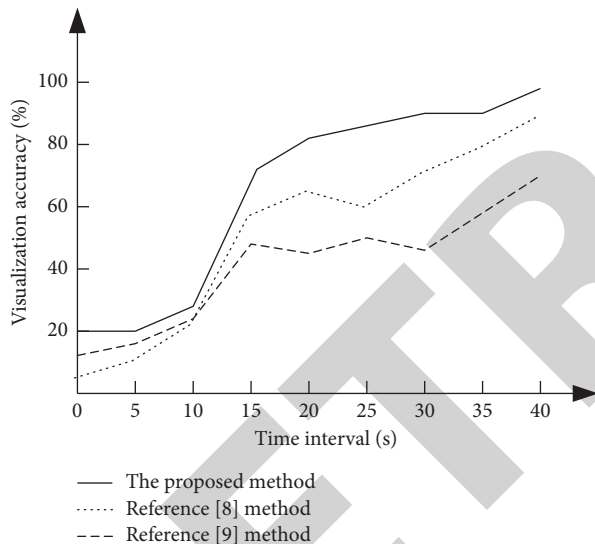


FIGURE 9: Comparison of visualization accuracy.

reference [9], the proposed method has shorter visualization time of unstructured data in social networks and can effectively improve the visualization efficiency of unstructured data in social networks.

**4.5. Accuracy Analysis.** On this basis, the visualization accuracy of social network unstructured data of the proposed method is further verified, and the visualization accuracy of social network unstructured data is taken as the evaluation index. The higher the visualization accuracy of social network unstructured data, the better the performance of social network unstructured data visualization model. By comparing the methods of reference [8], reference [9], and the proposed methods, the visualization accuracy of unstructured data of social networks with different methods is obtained, and the comparison results are shown in Figure 9.

It can be seen from the analysis of Figure 9 that the visual data accuracy of the three methods increases with the increase in time interval, and the accuracy of the three methods is significantly improved between 15s and 20s. This is because when the network starts running, there are a large number of slice data mixed and transmitted in the sub-network. Therefore, the visualization node will have a long time interval when receiving all the experimental data. When the visualization node collects all the mixed data, the visualization accuracy will be improved rapidly. After 15s, the accuracy of the visualization method of this design improves rapidly, and the accuracy gap between the comparison method and the method of this design is becoming larger and larger. Through comparison, it can be seen that the maximum difference between the comparison method and the design method is about 60%, and the minimum difference is about 20%. At the same time, the visualization accuracy of the comparison method is lower than that of the design method in the visualization process.

## 5. Conclusion

This study puts forward the dynamic simulation model of unstructured data visualization in social network, gives full play to the advantages of data visualization technology, and realizes the visualization of unstructured data in social network by combining acquisition fusion algorithm and visual feature processing algorithm on the Hadoop distributed computing platform. Its social network unstructured data visualization effect is good, which can effectively improve the stability and efficiency of social network unstructured data visualization. However, in the process of unstructured data visualization of social networks, data visualization technology is diverse and complex because it is an interdisciplinary subject and involves many research fields. For data in different fields, the requirements and standards of data visualization technology are inconsistent. Even for data in the same field, the requirements of visualization technology will have many branches. Therefore, in the next research, a unified data visualization standard needs to be formulated to effectively standardize the data information display and distribution results.

## Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

## Conflicts of Interest

The author declares no conflicts of interest regarding this work.

## Acknowledgments

This work was supported by the Research of E-Commerce Information Communication Mode on the Basis of Online Consumer Behavior and the Scientific Research Project from Education Office of Sichuan Province (No. Y201326813).

## References

- [1] G. Boella, A. Calafiore, E. Grassi, A. Rapp, L. Sanasi, and C. Schifanella, "FirstLife: combining social networking and VGI to create an urban coordination and collaboration platform," *IEEE Access*, vol. 7, pp. 63230–63246, 2019.
- [2] M. Naeem, "Do social networking platforms promote service quality and purchase intention of customers of service-providing organizations?" *The Journal of Management Development*, vol. 38, no. 7, pp. 561–581, 2019.
- [3] S. Sengupta and A. Vaish, "Social networking mood recognition algorithm for conflict detection and management of Indian educational institutions," *Social Network Analysis and Mining*, vol. 10, no. 1, Article ID s13278, 2020.
- [4] D. Wu, S. Fan, and F. Yuan, "Research on pathways of expert finding on academic social networking sites," *Information Processing & Management*, vol. 58, no. 2, Article ID 102475, 2021.
- [5] X. Qin, Y. Luo, N. Tang, and G. Li, "Making data visualization more efficient and effective: a survey," *The VLDB Journal*, vol. 29, no. 1, pp. 93–117, 2020.
- [6] J. C. Liu, C. T. Yang, Y. W. Chan, E. Kristiani, and W. J. Jiang, "Cyberattack detection model using deep learning in a network log system with data visualization," *The Journal of Supercomputing*, vol. 77, no. 10, pp. 1–20, 2021.
- [7] E. Qi, X. Yang, and Z. Wang, "Data mining and visualization of data-driven news in the era of big data," *Cluster Computing*, vol. 22, no. S4, pp. 10333–10346, 2019.
- [8] J. L. Kline and P. L. Volegov, "Toward 3D data visualization using virtual reality tools," *Review of Scientific Instruments*, vol. 92, no. 3, Article ID 033528, 2021.
- [9] L. Mrcic, S. Zajec, and R. Kopal, "Appliance of social network analysis and data visualization techniques in analysis of information propagation," *Intelligent Information and Database Systems*, vol. 7, no. 3, pp. 131–143, 2019.
- [10] C. Kavitha and X. Anita, "Task failure resilience technique for improving the performance of MapReduce in Hadoop," *ETRI Journal*, vol. 42, no. 5, pp. 748–760, 2020.
- [11] A. Mostafaiepour, A. Jahangard Rafsanjani, M. Ahmadi, and J. Arockia Dhanraj, "Investigating the performance of Hadoop and Spark platforms on machine learning algorithms," *The Journal of Supercomputing*, vol. 77, no. 2, pp. 1273–1300, 2021.
- [12] M. T. Wu, G. Srivastava, M. Wei, U. Yun, and C. W. Lin, "Fuzzy high-utility pattern mining in parallel and distributed hadoop framework," *Information Sciences*, vol. 553, pp. 31–48, 2020.
- [13] M. Maghsoudloo and N. Khoshavi, "Elastic HDFS: interconnected distributed architecture for availability-scalability enhancement of large-scale cloud storages," *The Journal of Supercomputing*, vol. 76, no. 1, pp. 174–203, 2019.
- [14] S. Zhang, "User interest based teaching information resource management method in HDFS mode," *Modern Electronics Technique*, vol. 42, no. 11, pp. 87–89, 2019.
- [15] K. Manjula and S. M. Sundaram, "Optimized approach (SPCA) for load balancing in distributed HDFS cluster," *SN Computer Science*, vol. 1, no. 2, Article ID s42979, 2020.
- [16] E. Azhir, N. Jafari Navimipour, M. Hosseinzadeh, A. Sharifi, and A. Darwesh, "A technique for parallel query optimization using MapReduce framework and a semantic-based clustering method," *PeerJ Computer Science*, vol. 7, p. e580, 2021.
- [17] M. K. Oo and M. A. Khine, "Topic extraction of crawled documents collection using correlated topic model in mapreduce framework," *Natural Language Computing*, vol. 6, no. 1, Article ID 01669, 2020.
- [18] P. Sowkuntla and P. Prasad, "MapReduce based parallel fuzzy-rough attribute reduction using discernibility matrix," *Applied Intelligence*, vol. 24, no. 4, pp. 1–20, 2021.
- [19] M. E. Porter, M. C. Hill, T. Harris, A. Brookfield, and X. Li, "The DiscoverFramework freeware toolkit for multivariate spatio-temporal environmental data visualization and evaluation," *Environmental Modelling & Software*, vol. 143, Article ID 105104, 2021.
- [20] Y. Liu, A. Béliveau, H. Besche et al., "Bayesian mixed effects model and data visualization for understanding item response time and response order in open online assessment," *Frontiers in Education*, vol. 5, Article ID 607260, 2021.
- [21] S. Rubab, J. Tang, and Y. Wu, "Examining interaction techniques in data visualization authoring tools from the perspective of goals and human cognition: a survey," *Journal of Visualization*, vol. 24, no. 1, pp. 397–418, 2021.
- [22] N. Blonder, B. C. Orsburn, J. Blonder, and C. A. Gonzalez, "Visual mass-spec share (vMS-Share): a new public web-based mass spectrometry visualization and data mining repository," *Journal of Proteomics & Bioinformatics*, vol. 12, no. 2, pp. 43–47, 2019.
- [23] C. H. Yu, H. Trier, and M. Slama, "A data mining and data visualization approach to examine the interrelationships between life satisfaction, secularization and religiosity," *Journal of Religion and Health*, vol. 58, no. 1, pp. 271–288, 2019.
- [24] M. Feng, J. Zheng, J. Ren et al., "Big data analytics and mining for effective visualization and trends forecasting of crime data," *IEEE Access*, vol. 7, no. 99, pp. 106111–106123, 2019.
- [25] T. Roughgarden and C. Seshadhri, "Distribution-free models of social networks," *Beyond the Worst-Case Analysis of Algorithms*, vol. 17, no. 12, Article ID 15743, 2020.
- [26] R. Rigia, M. Jalali, and M. H. Moattar, "A new method for community detection in social networks based on message distribution," vol. 25, no. 11, Article ID 12685, 2020.
- [27] A. Duarte-López, M. Pérez-Casany, and J. Valero, "The Zipf-Poisson-stopped-sum distribution with an application for modeling the degree sequence of social networks," *Computational Statistics & Data Analysis*, vol. 143, Article ID 106838, 2020.
- [28] L. I. Zuoyong, X. Wei, and J. Wang, "A univariate linear regression prediction model for environmental systems based on normal transform dimension reduction and error correction," *Acta Scientiae Circumstantiae*, vol. 39, no. 7, pp. 2455–2466, 2019.
- [29] G. Hesamian and M. G. Akbari, "A fuzzy nonlinear univariate regression model with exact predictors and fuzzy responses," *Soft Computing*, vol. 25, no. 4, pp. 1–16, 2021.
- [30] Z. Zhang and Z. Zhang, "Wireless network weak signal automatic detection simulation in chaotic background," *Computer Simulation*, vol. 36, no. 1, pp. 466–469, 2019.
- [31] L. Gu, H. Y. Lee, Y. Cao, and M. Liu, "Threshold adjustment compensation of asymmetrical optical noise," Patent 10454580, 2019.
- [32] Y.-Y. Won, S. M. Yoon, and D. Seo, "Ambient LED light noise reduction using adaptive differential equalization in Li-fi wireless link," *Sensors*, vol. 21, no. 4, p. 1060, 2021.
- [33] B. Chen, L. Zhang, and Z. Wu, "General iterative receiver design for enhanced reliability in multi-carrier differential chaos shift keying systems," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7824–7839, 2019.

- [34] B. Arvinti, A. Isar, and M. Costache, "Effective wavelet algorithm for an automatic detection of atrial fibrillation," *Advances in Electrical and Computer Engineering*, vol. 20, no. 3, pp. 49–56, 2020.
- [35] W. Li, J. Chen, J. Li, and K. Xia, "Derivative and enhanced discrete analytic wavelet algorithm for rolling bearing fault diagnosis," *Microprocessors and Microsystems*, vol. 82, Article ID 103872, 2021.
- [36] L. d. S. Santos, A. R. Secchi, and D. M. Prata, "An adaptive sequential wavelet-based algorithm developed for dynamic optimization problems," *Computers & Chemical Engineering*, vol. 121, pp. 465–482, 2019.

RETRACTED