Data Fusion Method of Data News Report Based on Relative Difference Coefficient

Yang Yang

College of Humanities, Geely University of China,Chengdu,641423,China
E-mail: bagbear1111@hotmail.com

Abstract—The traditional data fusion methods of data news reports have the problems of poor data fusion accuracy and poor data fusion effect. A data fusion method of data news reports based on relative difference coefficient is proposed. By analyzing the functional requirements of data fusion, collect news report data, preprocess the collected data based on the relative difference coefficient, and fuse the data and news report data according to the processing results. Comparative experiments are used to verify that the data fusion accuracy of this method is higher, the data fusion effect is better, and it is of great promotion value.

Key words: relative difference coefficient; data journalism; news reports; data fusion; method;

I INTRODUCTION

"Data news" also called "data-driven news", the current academic circles on the has not yet formed a unified definition, but its core connotation, analysis of the data mining, filter and visualization, thus data as expression form of news reports, so the data analysis technology to become an integral part of the news production process [1]. It through the study of the integration of data, realizing the news of digging, the data used in news is public, the tool resources are shared, it is a trait data news." Peng Lan described the future development prospects of data journalism from a technical perspective, with the support of big data and other technologies, information resources needed in news will be collected more and more automatically, filtered and analyzed through relevant technologies. In fact, data journalism is nothing new. As early as 1901, the guardian attempted to present the battle situation of the British empire during the second Boer war by visualizing data [2-4]. However, it was not until 2009 that data journalism emerged and spread widely as a new word, showing its difference from traditional news reporting, to put it simply, data journalism pursues to show the news facts behind abstract data through visual means such as data and charts, supplemented by a small amount of text. In terms of practical operation, different from traditional news reporting, data journalists need to have enough awareness of data, explore news value from interesting data, analyze and process data, and finally provide systematic, objective and comprehensive reports for readers in the easiest way to understand by visualization technology.

Research on the presentation mode of data news. At present, domestic researches on this issue only focus on a few papers, such as brief analysis of narrative innovation in data news reporting, Visualization of "data news", new transformation of the role of journalists, and preliminary study of visual narrative of data news [5]. Summarizes such data journalism handbook presents several forms, using data and graphics to convey to the readers and effective news and information. For example, Hans Rosling used Gapminder software to visual report on global poverty, obtained from the attention of millions of people around the world. David mike's grace through refining the core data, visualizations of fiscal expenditures and the pollution caused and prevented by Iceland's volcanic eruption were presented to the public. Interaction with readers can be realized through open data to show the transparency of the news production process. For example, "data blog" of the guardian [6] is a successful case of sharing data and news with readers in the form of information ICONS, maps and interactive renderings. Based on data, data analysis software is used to process massive information, such as the investigative report on mps by daily telegraph [7]. Explaining the connection between news messages and individuals, such as the interactive reports on the budget produced by the BBC and the financial times [8], conveying useful information while explaining how it is relevant to citizens' lives. Data journalism has become a major innovation in news reporting in the era of big data. In recent years, the New York times [9], the guardian, BBC and other world-renowned media have started data journalism business. And the Las Vegas sun's series on medical "do no harm". The concept of data journalism has been popping up in seminars and academic writings since the associated press's textual analysis of the Iraq war records made it even more prominent. Our portals such as sina, sohu, netease, tencent also began to make a difference, respectively opened "graphic", "number", "read the column", "data control" and other columns, try to present the facts to data. But basically is given priority to with digital+ images, there is a big gap compared with the European and American countries. At present, the only book that can be collected specifically to discuss "data journalism" is "data journalism handbook". Championed by the European journalism centre and the open knowledge foundation, it was compiled by industry and academia from BBC, guardian, financial times, New York times and other media focusing on the field of data journalism through network collaboration [10]. On the one hand, the book explains what data journalism is and the significance of data for journalists. On the other hand, it describes the production process of data journalism from three aspects of data acquisition, data understanding and data transmission by using cases of international mainstream media operating data journalism. Therefore, the data fusion of data news reports has become the focus of current research. Based on this, this paper proposes a data fusion method of data news reports based on relative difference coefficient.

II DESIGN OF DATA FUSION METHOD FOR DATA NEWS REPORT BASED ON RELATIVE DIFFERENCE COEFFICIENT

A, Analyzing the functional requirements of data fusion

The problems of traditional data fusion data fusion

methods are as follow. Facing huge amounts of data, the traditional centralized data storage is too simple, all data are piling up in together, easy to cause the loss of performance. At the same time, the data of a single storage requirement for servers and other hardware facilities is higher, anti-destroying ability is bad, if the server because of some fault collapse is likely to cause loss of large amounts of data. The traditional data fusion method consumes a lot of time due to its own performance problems in preprocessing operations such as formatting and removing redundancy of a large number of original data collected from sensor networks, resulting in low efficiency [11-13]. Traditional data fusion methods for large-scale data sets in computing and storage, due to their own data fusion method of computing power and storage capacity constraints, and the algorithm of parallel implementation is poorer, leading to low efficiency of the data fusion method, anti-destroying ability is poor, inflexible, not suitable for current application scenario for the era of big data. When the traditional data fusion application algorithm realizes the specific data fusion process, the centralized data processing does not make good use of some parallel characteristics of the algorithm, and the flexibility is poor. And implements the improved data fusion algorithm by using map reduce programming model in parallel. Through the HDFS and HBase high throughput, greatly accelerating the speed of data reading and data storage, the use of HDFS and reliable data storage capacity, effectively enhancing the capacity of the cluster data fusion method of invulnerability, and robustness, using them to the characteristic of high throughput, the data fusion task parallelization, enhancing the ductility of the data fusion method and scalability, and improving the efficiency of data fusion [14]. Through the effective combination of various technologies, the whole data fusion methods can well meet the actual application scenarios. Data storage modules have two major requirements. One part is HDFS storage and the other part is HBase storage. HDFS storage requirements are mainly for convenient storage of data sets to be fused, including basic functions such as viewing directories, creating folders, moving or renaming files or folders, deleting files or folders, and uploading and downloading files. HBase storage is used to store data sets to be fused in HBase tables to facilitate subsequent data acquisition and data fusion. It mainly includes the addition and deletion of data tables and the addition and deletion of data units in data tables. Data fusion module includes three requirements, selecting data to be fused, core computing of data fusion, and display of fusion results. The requirement of data to be fused is selected to provide users with an optional list of data to be fused and guide them to select a specially fused data set [15]. The core computing requirement of data fusion is to take the data set selected by the user as the input and get the final data fusion result by calling the algorithm flow of data fusion method. The requirement of fusion result display is to display the fusion result obtained by data fusion through visual means, so that users can view and analyze the data fusion result more intuitively. User management module, this module in addition to the most basic login registration requirements, but also for different users to do different permissions. There are two groups: administrator group and common user group. Thus, the data fusion architecture is shown in figure 1 below.
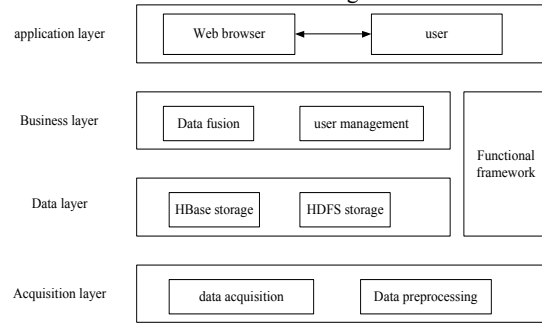

Fig. 1 Data fusion architecture

As shown in figure 1, the data fusion method based on the relative difference coefficient consists of HDFS, HBase, front-end page and back-end server, which are distributed file storage data fusion methods of Hadoop platform. It is mainly divided into four functional modules: data acquisition and preprocessing module, data storage module, data fusion module and user management module. The data fusion method adopts B/S structure, that is, browser or server structure. After a successful login, users can interact with the data fusion method through the front-end page of the browser. The application layer mainly refers to the user and the browser. Users operate the browser to perform corresponding operations on the back-end data fusion module and the storage module of the data layer. The browser is also responsible for displaying the operation results to the user. The business layer includes the core module of the data fusion method, the data fusion module, and the essential user management module. The data used by the data fusion module needs to be obtained from the storage module. The user management module provides different permissions for different users to ensure the security of the data fusion method. The acquisition layer includes data acquisition and data preprocessing functions. Wireless sensor network is used to collect the characteristics of the monitored target and carry out corresponding preprocessing for subsequent use.

B, Gathering news report data

Data in the big data era mainly refers to unstructured data in many forms, including text, audio, video, animation, pictures and so on. These large-scale data are the basis of data news reporting. From the practice of the employment sector, the second-hand data obtained from the public databases of the government, enterprises and social organizations and the first-hand data captured by the user survey conducted by the media are the main data sources available to the news media [16-17]. The requirements for data authenticity and reliability make the rise of open data and data "crowdsourcing" possible. An open Internet environment is the fundamental premise for the rise of data news. On the one hand, data openness is related to the ability of journalists to process information and dig news clues, and help the audience accurately understand the ongoing news events and their impact on personal life; On the other hand, it is also related to the audience's evaluation of news media and cognition of social environment. The government is the largest owner of large-scale information and an important source of

1035

public data, so it should do more in data opening.

In the era of big data, network platforms or other digital platforms have become platforms for data collection, rather than just publishing news mechanically. At present, the completion of data collection through crowdsourcing is mainly reflected in three aspects: first, professional media can publicly collect text, pictures, animation, video and other forms of materials for a topic that the public may be interested in. The audience can submit them online, and then the reporter can complete the news compilation through the data analysis background. The audience can also actively provide information sources or express opinions on ongoing events on the open platform of the media. Secondly, the professional media will launch the interface of information chart, open the data entrance for the public to query at any time, and the formed data will be used by the media. Third, other platforms such as online communities and handheld mobile devices are also important data sources. Registered users are encouraged to download a mobile terminal app, which can publish what they have seen and heard, comment, forward and participate in discussion [18].

The news report data are displayed in the form of cluster centers, in which m D-dimensional cluster centers are the positions of particles, the corresponding speed is $m \times D$, and there is a fitness. The classification of news report data samples is established by using the nearest neighbor criterion, and the clustering division of news report data samples is obtained, which is expressed as: $A = \{A_1, A_2, \text{L}, A_m\}$. at this time, the particle fitness is defined as:

$$f = \frac{hJ_A}{U} \tag{1}$$

$$J_A = \sum_{k=1}^{m} \sum_{X_i \in A_k} d(X_i, C_k)$$

In the formula, and k describe the category, J_A describes the nearest neighbor criterion, C_k describes the clustering center of class k, X_i describes the particle fitness, m describes the number, i and j describe the coefficient, C_m^2 describes the combination number, h describes the constant, the value is 20, $d(X_i, C_k)$ describes the distance between X_i and C_k, and the Euclidean distance at this time is $d(X_i, C_k) = \|X_i - C_k\|$.

Today's journalists should help their audiences make sense of information, but this does not mean simply adding explanation or analysis to a story. Instead, they must verify the reliability of the information and then organize it so that it can be quickly and effectively understood. This means that data is just a tool. Analyzing and screening data, mining stories worth telling, and helping audiences establish the connection between data and individuals are the solutions that news media should take in the face of changes in information environment. Data analysis is not entirely dependent on technical judgment, talent is the core of data journalism. As Professor Fu Yuhui of communication university of China said, "even though new technology can cause new fission and differentiation of news business types, such fission and differentiation still cannot completely replace the core position of human in professional news content production by the application of machine."

C, Data processing based on relative difference coefficient

The data fusion system based on relative difference coefficient takes distributed storage and map reduce programming model as the core, preprocesses and stores the original data collected by wireless sensor network in the distributed system, and uses each node in the cluster to process the whole data fusion task in parallel to improve the system performance. In a data fusion system, upper-layer applications deliver data fusion and storage tasks layer by layer by invoking interfaces provided by lower-layer services. News data includes a lot of unstructured data, including text data, audio data, video data and so on. As audio and video data are only archived as historical data, they will not be converted and stored here. This paper aims at structured transformation of text data in news data. Whole life cycle of the emu data integration management in the face of the text data is divided into two cases, one is technical documents, technical parameters such as the emu resume information, another is due to business data interface requirements, such as export in train-ground communication system history state data can only be in the Excel spreadsheet form, therefore this paper to obtain such data for text data. The data processing process is shown in figure 2 below.
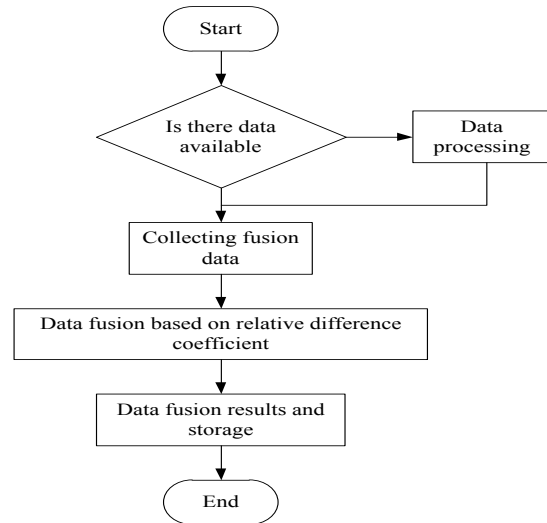

Fig. 2 Data processing process

As shown in figure 2, in the process of data processing, the data to be fused is firstly analyzed whether it is available, then the data to be fused is selected, and the relative difference coefficient is used to fuse the data, and finally the fusion results are stored. The Main class inherits thread class and is the starting point of the whole data collection and preprocessing task. Portid represents the port number of the main sensor node, _mac

represents the unique identifier of the node, _old box num represents the old box number, _new box num represents the new box number. The connection is checked for readiness through the prepare communication () method, and the collect thread is called through the run() method for data collection. The collect thread class implements the runnable interface, which is responsible for sending data collection instructions and collecting data to the primary node of the sensor network. Data_from_serislport is used to store the collected data. Data_num_from_serialport indicates the number of collected data. The handle thread class is responsible for processing the data retrieved by the collect thread by calling the corresponding method in handle data. Handle data parses the collected original data, formats the original data, converts it to the required data format for subsequent processing, removes unnecessary redundant data, stores the formatted data to the historical database, processes the data to be fused and stores it in the HBase table for future use. Data_to_check indicates the data to be parsed, portid indicates the port number, data_num_to_check indicates the amount of data to be processed, get date () is used to get the system timestamp, check() is used to parse raw data and remove redundancy, The parsed data is stored in the historical database, and the data to be merged is processed and stored in HBase tables.

In order to ensure the authenticity of data, this paper conducts regression analysis on data news as follows:

$$y_i = \beta_{i0} + \sum_{k=1}^{n} \beta_{ik}(u_i, v_i)x_{ik} + \varepsilon_i \quad i = 1, 2, ..., n \ (1)$$

In formula (1), y_i is regression variable, β_{i0} is the initial data of the i th sample, (u_i, v_i) is the k th data of the i th sample. (u_i, v_i) is the regression coordinate of the i th sample, x_{ik} is the k th variable of the regression equation. ε_i as the residual error, n for the constant. Assuming an independent normal distribution, so that:

$$\varepsilon_i = N(0, \sigma^2) \tag{2}$$

$$Min\sum_{i=1}^{n}\beta_{ik} = \frac{y_i - \beta_{i0}}{\frac{\sigma}{\sum_{k=1}^{n} x_{ik}\varepsilon_i}} \tag{3}$$

In formula (2) and formula (3), σ is the relative difference coefficient of news data, N is regression parameter, $Min\sum_{i=1}^{n}\beta_{ik}$ is the regression value of data. The regression indexes of news data are as follows:

$$\hat{\beta_i} = Wx_{ik} \cdot Wy_i \tag{4}$$

$$\hat{y_i} = x_i \hat{\beta_i} \tag{5}$$

In formula (4) and formula (5), $\hat{\beta_i}$ is regression correlation coefficient, W is the difference index, $\hat{y_i}$

is the regression index of news data [19-20]. According to this method, the effect of data fusion can be improved and the authenticity of news data can be guaranteed.

D, Realizing data fusion of news reports

The data fusion method based on Kalman filtering, data fusion method based on Bayesian, news reports data fusion based on relative difference coefficient method, the data fusion method based on production rule, the data fusion method based on neural network and data fusion method based on fuzzy set theory. The latest research results of each method are introduced. Then, this paper expounds the basic idea of the data fusion algorithm based on the relative difference coefficient, and analyzes the shortcomings of the algorithm in detail. Then, starting from solving the problem of evidence conflict, this paper proposes a data fusion algorithm based on the improved relative difference coefficient. By introducing The Babbitt distance and making some special processing in combination with the evidence theory scene, the evidence confidence level was obtained, and the evidence BPA was adjusted according to the evidence confidence level.

Since traditional data fusion methods have insufficient computing power and poor parallelization when facing large-scale data, this paper also combines data fusion technology with big data technology to realize data fusion method based on Hadoop platform. This paper studies and analyzes the requirements of data fusion method, basic structure and method basic processing process, and then combines it with the characteristics of Hadoop platform to implement the improved data algorithm in this paper, completes the design and implementation of four basic functional requirement modules. Finally, through the experimental demonstration on Hadoop platform, The feasibility of the algorithm and the availability and reliability of the data fusion method are verified.

In the future work, the author will further study the proposed data fusion algorithm to make the improved data fusion algorithm suitable for most application scenarios. In view of the four shortcomings of the relative difference coefficient, the future research will strive to solve more problems, so that the applicability and practicability of the data fusion algorithm is more powerful. Combined with advanced technology, it is intended to make the whole data fusion method more intelligent and automatic, so that the data fusion method can better serve users, and the goal is to bring users the ultimate user experience. In general, telling news stories in a visual way is the essential connotation of data journalism. Data itself does not generate value, and the core is to use data to convey the understanding of social significance. Video, photos, data fusion, tables and other visual means are the fastest and most effective way to convey information. It is a professional quality that data journalists must possess to visually present data analysis results and help audiences understand deeper meanings while attracting their attention. Compared with lengthy text expression, data fusion, animation, video and other multimedia means have unique advantages in conveying information. From the practice of media at home and abroad, data fusion, as a way of visualization, is playing an increasingly important role. The New York times, The guardian and other internationally renowned media all

have visualized data news teams, and the graphic design also emphasizes interactivity to satisfy the audience's pleasure of "discovery". Data fusion provides a new perspective for the audience to understand the world, which mainly presents three development directions. Firstly, the advantages of interactive charts and dynamic charts over static charts are increasingly prominent, and the audience is more and more involved in the production process of data fusion. Audiences can just drag the mouse to get the information they want, instead of being pushed by the news media on their own. Users can also get corresponding information by inputting data to improve the quality of decision-making. The output results of the chart will change with the input data, and the sharing button will be designed to allow them to share to the social network platform with one click while browsing. This strategy combining personalization and socialization will be an important direction for the development of data news in the future. Secondly, data maps, time lines, interactive charts and other elements will be integrated to jointly tell complex news events from multiple dimensions, especially in sorting out dynamic development process, describing the relationship between characters, and showing the overall development of things. The organic combination of text, map, video, animation and audio to achieve the coordination of visual and auditory senses can better attract the audience's attention and greatly improve the readability of the report. Thirdly, the availability of large amounts of unstructured data has made it an essential skill for data journalists to be proficient in visualization software, such as Google fusion tables, tableau public, and Google's linear and pie chart tools. Data journalism will be journalism and statistics, design, programming and other disciplines together, the intersection of computer technology and traditional editing ability will be more and more, can be said to truly reflect the "collection of science, technology and art". Only by combining liberal arts with science and engineering and improving the ability of journalists to use data can we bring qualitative breakthrough for the development of data journalism in the future.

III THE EXPERIMENTAL TEST

A, The experiment to prepare

In order to improve its processing speed, the experiment independently builds a Hadoop cluster environment, pre-installed Centos operating system, Hadoop, Hive and IDE tools and so on, and uses 6 X86 PCS as fusion devices. One serves as Master and Name Node, and five serves as Slave and Data Node. The hardware and software configurations of the six computers are the same, as shown in table 1.

Table 1 The experimental environment

Cluster environment	The operating environment
X 86 frame built PC	Operating system: Centos 7
CPU: 4 core 2.5 GHz	Hadoop version: Hadoop-2.3.0
Memory: 8 GB	Hive version:Hive-0.12.0
Hard disk: 1TB	Java development environment:JDK1.8.0.25
Nic: gigabit ethernet nic	ETL tools: Pentaho (Kettle) v7.0

As shown in table 1, in this cluster environment and operating environment, data is collected and preprocessed first, and related parameters of the processed data are shown in table 2 below.

Table 2 Related parameters after data news processing

The field name	Fields instruction
Data news type	Data news type
News batch number	News batches.
Total	The total number
Acquisition speed	Gathering speed

As shown in table 2, the parameters after data processing are represented by the above fields. In addition, progress index is used to calculate the accuracy of data fusion, and the formula is as follows:

$$precision = \frac{C}{C+I} \qquad (6)$$

In formula (6), $precision$ is the accuracy of data fusion, C is the amount of attribute data for precise data fusion, I represents the amount of data that cannot be accurately fused.

B, The experimental results

The traditional method was compared with the method designed in this paper to verify the accuracy of data fusion of the two methods, as shown in table 3 below.

Table 3 The experimental results

The amount of data	Data fusion accuracy of traditional methods /%	The method designed in this paper is accurate in data fusion /%
1000	94.2%	99.9%
2000	93.6%	99.9%
3000	92.4%	99.8%
4000	91.2%	99.8%

As shown in table 3, the accuracy of data fusion in traditional methods is poor, both below 95%, which is not in line with the data fusion background. However, the precision of data fusion in the method designed in this paper is high, approaching 100% infinitely, which is in line with the research purpose of this paper.

In order to further verify the effectiveness of this method, the traditional method and this method are used to calculate the relative difference coefficient and verify the calculation accuracy of the two methods. The higher the calculation accuracy, the higher the data fusion accuracy of data news reports and the better the fusion effect. The comparison results are shown in figure 3.
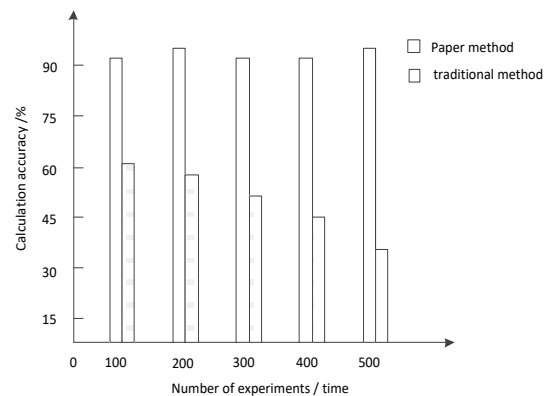

Fig. 3 Comparison of calculation accuracy of relative difference coefficient between the two methods

According to figure 3, the calculation accuracy of

the relative difference coefficient of this method is more than 90%, which is higher than that of the traditional method, indicating that the data news report data fusion accuracy of this method is high and the fusion effect is good.

IV CONCLUSION

This paper mainly discusses from three aspects. Firstly, it systematically combs the evolution history and development status of data journalism, and discusses in detail the characteristics of "data journalism" compared with traditional news in terms of news production mode, presentation form and reporting strategy under the background of big data. Secondly, combined with classic cases at home and abroad, the practice of domestic and foreign media in the field of data journalism is compared from the perspective of visual presentation form, and several domestic development limitations are put forward. Thirdly, it discusses the future development trend of data journalism and the new requirements and missions for journalists from three aspects of data acquisition, data analysis and data presentation. The rise of data journalism transformation provides a new way for traditional media and may, but the new report way in practice still faces problems and deficiencies. Firstly, excessive emphasis on technical factors and ignores the subject status of people could lead to journalists in the abyss of huge amounts of data, thus weakening the field to cover the ability. Journalists should not blindly follow data, but filter and analyze it based on value judgments. In addition to the second-hand data captured from the government, enterprises, social organizations and other public databases, journalists must enhance the awareness of the field, through on-the-spot interviews to obtain first-hand data. To have the ability to identify false, meaningless and privacy infringing data, extract real and valuable content from them. Can discover the essence of the problem from the abstract data, reveal the correlation between the data and the law of development. As Franks says, "The key to successful analytics is not the tools and techniques themselves, but the people who use them are the core elements of success." Secondly, the limited data resources mastered by news media greatly affect the development pace of data journalism. It is embodied in the following aspects. Firstly, government-led public database is the most important data resource for news media, but the most prominent problem at present is that China lags far behind the western countries in data opening. It didn't enact the regulations on the disclosure of government information until around 2008, and the lack of judicial binding force and social supervision leads to poor implementation effects. The second is, the news media big data obtained mainly from network data, represented by social media data and IOT sensor data, existing data for digital processing of traditional media, government, enterprises, institutions and other departments of open data, academic data reuse, media to survey to collect data, and so on six aspects.

The data news report data fusion method based on relative difference coefficient can improve the production efficiency and work quality of data news report. Therefore, in the next research, in order to improve the timeliness of data news report, we will further discuss the data news fusion method of multi-channel and multi-party method.

REFERENCES

[1]   Wang K , Hossain S , Habib K N (2021). A hybrid data fusion methodology for household travel surveys to reduce proxy biases and under-representation of specific sub-group of population. Transportation,(1):1-36.

[2]   Ren C , Liang Y J , Lu X J , et al(2019). Research on the soil moisture sliding estimation method using the LS-SVM based on multi-satellite fusion. International journal of remote sensing, 40(5-6):2104-2119.

[3]   Wang Y , Cong J , Wang P , et al(2019). A Data-Fusion Approach for Speed Estimation and Location Calibration of a Metro Train Based on Low-Cost Sensors in Smartphones. IEEE Sensors Journal, PP(99):1-1.

[4]   Xue H , Yu M , Chen C (2019). Research on Novel Correlation Coefficient of Neutrosophic Cubic Sets and Its Applications. Mathematical Problems in Engineering, 2019:1-10.

[5]   Hashim H A , Eltoukhy A (2021). Landmark and IMU Data Fusion: Systematic Convergence Geometric Nonlinear Observer for SLAM and Velocity Bias. IEEE Transactions on Intelligent Transportation Systems:1-10.

[6]   Xue T , Zheng Y , Geng G , et al(2020). Estimating Spatiotemporal Variation in Ambient Ozone Exposure during 2013–2017 Using a Data-Fusion Model. Environmental Science and Technology, 54(23):14877-14888.

[7]   Li K , Liu Y , Wang S , et al(2021). Multifidelity Data Fusion Based on Gradient-Enhanced Surrogate Modeling Method. Journal of Mechanical Design, 143(12):1-24.

[8]   Wang Z , Yu G , Zhou B , et al(2021). A Train Positioning Method Based-On Vision and Millimeter-Wave Radar Data Fusion. IEEE Transactions on Intelligent Transportation Systems, PP(99):1-11.

[9]   Cai X , Zhang C , Yang Y , et al(2020). Data Fusion Method of Measurement Lag Compensation for Multi-rate MIMU/FOG/GNSS Compound Navigation. IEEE Sensors Journal, PP(99):1-1.

[10] Xian T , Wu J , X Zhang(2020). Study on the operating characteristics of a solar heat pump water heater based on data fusion. Solar Energy, 212:113-124.

[11] Wang M , Zhang Z , Xie Y , et al(2020). Fault diagnosis analysis and health management of thermal performance of multi-source data fusion equipment based on fog computing model. Thermal Science:318-318.

[12] Tan C , Sun Y , Li G , et al(2020). Research on gesture recognition of smart data fusion features in the IoT. Neural Computing and Applications, 32(22):16917-16929.

[13] Fca B , Phh C (2020). Prediction of human odour assessments based on hedonic tone method using instrument measurements and multi-sensor data fusion integrated neural networks. Biosystems Engineering, 200:272-283.

[14] Bai J , Li S , Huang L , et al(2021). Robust Detection and Tracking Method for Moving Object Based on Radar and Camera Data Fusion. IEEE Sensors Journal, PP(99):1-1.

[15] Zuo J , Meng L , Li C , et al(2020). Construction of community life circle database based on high-resolution remote sensing technology and multi-source data fusion. European Journal of Remote Sensing, (3):1-16.

[16] F Wang, Du J , Zhao Y , et al. A Deep Learning Based Data Fusion Method for Degradation Modeling and Prognostics. IEEE Transactions on Reliability, 2020, 70(2):775-789.

[17] Schliep E M , Collins S M , Rojas-Salazar S , et al. Data fusion model for speciated nitrogen to identify environmental drivers and improve estimation of nitrogen in lakes. The Annals of Applied Statistics, 2020, 14(4):1651-1675.

[18] Ali, Karsaz. Chattering-free hybrid adaptive neuro-fuzzy inference system-particle swarm optimisation data fusion-based BG-level control. IET Systems Biology, 2020:31-38.

[19] Lamti H A , Khelifa M B , Hugel V . Cerebral and gaze data fusion for wheelchair navigation enhancement: case of distracted users. Robotica, 2019, 37(2):246-263.

[20] Wang X , Lu S , Chen K , et al. Bearing Fault Diagnosis of Switched Reluctance Motor in Electric Vehicle Powertrain via Multi-sensor Data Fusion. IEEE Transactions on Industrial Informatics, 2021, PP(99):1-1.