Retraction

Retraction: Research on Computer Information Processing Technology from the Perspective of the Big Data Era (J. Phys.: Conf. Ser. 2033 012147)

Published 9 September 2022

This article has been retracted by IOP Publishing following an allegation that raises concerns this article may have been created, manipulated, and/or sold by a commercial entity. In addition, IOP Publishing has seen no evidence that reliable peer review was conducted on this article, despite the clear standards expected of and communicated to conference organisers.

The authors of the article have been given opportunity to present evidence that they were the original and genuine creators of the work, however at the time of publication of this notice, IOP Publishing has not received any response. IOP Publishing has analysed the article and agrees there are enough indicators to cause serious doubts over the legitimacy of the work and agree this article should be retracted. The authors are encouraged to contact IOP Publishing Limited if they have any comments on this retraction.

Retraction published: 9 September 2022
Research on Computer Information Processing Technology from the Perspective of the Big Data Era

Jinmeng Li*
School of Management, Tianjin University of Technology, Tianjin, China

*Corresponding author: jinmengli2019@tjut.edu.cn

Abstract. With the development of China's social economy and science and technology, the computer age is also developing rapidly. Nowadays, the development and progress of society is inseparable from the application of computer network technology. At the same time, the development of Internet technology has enabled us to enter the big data society a long time ago. Information technology not only promotes the development of society, but also computer technology is very important to people's lives, so the number of computer users is very large at present. It is not easy to deal with big data due to its huge quantity, complex form and unstructured characteristics, which often requires analysts to have a high reserve of relevant knowledge such as statistics and computer technology.

Keywords: big data, data mining, computer, information processing technology.

1. Introduction
Big data is not only the huge amount of data, but more importantly, the analysis and processing of massive data, after which the useful information contained in these massive data can be obtained. At present, big data has been widely used in various fields such as social economy and science, so the method of big data analysis becomes particularly important. It can be said that the key factor determining the value of information is the use of analytical methods[1]. Basic knowledge about big data is shown in Figure 1.
Figure 1. The "4V" characteristics of big data.

Volume: Many factors lead to an increase in the amount of data. The growth of industrial engineering and manufacturing has made the data in the production process larger and larger, but computer technology has made data collection more and more convenient. Before, excessive data volume was a storage problem, but as storage costs were reduced, problems such as the relationship between big data and the analysis methods of the value behind big data began to appear.

Velocity: The velocity of data flow is very fast because so much data must be processed in time. RFID tags, sensors and smart meters have the ability to process large amounts of data in real time. For many companies, processing data is a big problem.

Variety: From a formal perspective of data organization, data types can be divided into structured data and unstructured data. Structured data is the data that can be stored in the database, such as bank transaction data, civil aviation flight information data, etc. While unstructured data is data that cannot be represented by predefined data models or stored in relational databases, such as office documents, images, E-mail, and videos, merging and managing the diversity of different data is something many organizations are still grappling with.

Value: Huge data will slow down the data analysis speed, and the variety of big data, the density of valuable data in these big data is very low. With big data technology, these valuable data can be used to predict and analyze future trends and patterns.

Computer information processing technology has the characteristics of automation, relevance, programming, etc., is an important branch of digital program innovation, it will also be helpful to the development of computer technology. With the help of big data, industrial engineering will be affected by computer technology. In addition, it will also affect many details in the production process. Therefore, we have high expectations for the development of computer technology, which is undoubtedly a great pressure for employees in the computer field. The computer information processing technology is closely related to the development of many industries. A large amount of data information will also be affected by information processing technology. No matter what the current situation of information technology is, the high requirements we put forward are challenges for it [2]. With the increasing amount of data, the types of information are gradually developing in a complex direction. In order to solve the above problems, relevant fields can only gradually improve the quality and efficiency of information data processing. It is imperative to update the information processing technology.

2. Partitioning around medoids algorithm

In this paper, partitioning around medoids (PAM) in cluster analysis is taken as an example to briefly explain the basic operation scheme of this algorithm for big data.

The specific algorithm process is as follows [3]:

\[\text{Algorithm: Partitioning around medoids (PAM)}\]

1. Initialization: Choose k medoids from the dataset.
2. Assignment: Assign each data point to the nearest medoid.
3. Update: For each medoid, recalculate its new position as the average of the data points assigned to it.
4. Repeat steps 2 and 3 until no changes occur.

This algorithm is particularly useful for large datasets and can be applied in various fields such as data mining, pattern recognition, and machine learning.
Step 1. Randomly select K points from n sample points in the population as medoids.
Step 2. According to the closest principle with medoids, allocate the remaining n-K points to the class represented by the best medoids.
Step 3. If other data points in class i are new medoids, then we will calculate the value of the standard function., iterate all possibilities, and select the point corresponding to the smallest criterion function as the new medoids.
Step 4. Repeat the above two steps. When all medoids no longer change or the calculation has reached the maximum number of iterations, the calculation stops.
Step 5. Outputs the finalized K classes.
Criterion function: minimizes absolute error

\[ E = \sum_{i=1}^{k} \sum_{p \in C_i} dist(p, o_i) \]  

(1)

Cluster the 10 data in Table 1,
Randomly pick k=2 center points: c_1=(3,4),c_2=(7,4). Then calculate the distance from all points to these two points, and the results are shown in Table 2. It can be seen that the black body is the distance value with the smaller distance to the two center points. Then according to Table 2, we can classify all data points:
Cluster1 = {(3,4)(2,6)(3,8)(4,7)}
Cluster2 = {(7,4)(6,2)(6,4)(7,3)(8,5)(7,6)}
The cost at this time is 20.

| Table 1. Raw data for cluster analysis. |
|----------------------------------------|
| X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 | X_8 | X_9 | X_10 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 2   | 3   | 3   | 4   | 6   | 6   | 7   | 8   | 7   | 5    |
| 6   | 4   | 8   | 7   | 2   | 4   | 3   | 4   | 5   | 6    |

| Table 2. Distance between two points. |
|--------------------------------------|
| Data object | \(X_i\) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | cost |
|-------------|--------|---|---|---|---|---|---|---|---|---|----|------|
| Distance to | \(c_1=(3,4)\) | 3 | 0 | 4 | 4 | 5 | 3 | 5 | 4 | 6 | 6 | 11   |
|             | \(c_2=(7,4)\) | 7 | 4 | 8 | 6 | 3 | 1 | 1 | 0 | 2 | 2 | 9    |

Pick a non-central point \(o'\), assuming that the selected point is \(X_7\), that is \(o'=(7,3)\). Then these two center points become \(c_1=(3,4)\) and \(o'=(7,3)\).

| Table 3. Cost after replacing the center point. |
|------------------------------------------------|
| i | c_1 | Data objects(\(X_i\)) | cost | o' | Data objects(\(X_i\)) | cost |
|---|-----|------------------------|------|----|------------------------|------|
| 1 | 3   | 4                      | 2    | 6  | 3                      | 7    | 3 | 2 | 6 | 8  |
| 2 | 3   | 4                      | 3    | 8  | 4                      | 7    | 3 | 3 | 8 | 9  |
| 3 | 3   | 4                      | 4    | 7  | 4                      | 7    | 3 | 4 | 7 | 7  |
| 4 | 3   | 4                      | 6    | 2  | 5                      | 7    | 3 | 6 | 2 | 2  |
| 5 | 3   | 4                      | 4    | 4  | 3                      | 7    | 3 | 6 | 4 | 2  |
| 6 | 3   | 4                      | 7    | 4  | 4                      | 7    | 3 | 7 | 4 | 1  |
| 7 | 3   | 4                      | 8    | 5  | 6                      | 7    | 3 | 8 | 5 | 3  |
| 8 | 3   | 4                      | 7    | 6  | 6                      | 7    | 3 | 7 | 6 | 3  |
As shown in Table 3, the total cost = 3+4+4+2+2+1+3+3 = 22, which is larger than the previous cost=20, so the cost of this replacement becomes larger and we finally do not make this replacement. This is just trying to replace c2 with X7, we should calculate all the points except for c1 and c2 for c1 and c2 respectively, and calculate the costs after these substitutions. If there is a cost less than 20 then we will use the center point corresponding to the minimum cost as the new center point. At this point, one iteration is complete. Repeat the iteration until convergence.

3. Computer information processing technology in the era of big data

3.1. The status quo of computer information processing technology in the context of big data

In recent years, with the popularization of the Internet of Things, big data and cloud computing, explosive growth has become a significant feature of data growth. Under the situation of increasingly fierce business competition, many enterprises adopt computer information processing technology to take advantage of the accuracy and speed of big data processing and the fast transmission speed to conduct in-depth excavation within the enterprise and the industrial chain. However, some people only look at the advantages and superficial aspects of big data, there is no extensive study of computer technology. The rapid growth of data leads to the occurrence of data abuse, which also shows that computer technology is seriously constrained in the big data environment. To solve this problem, relevant personnel should improve the advanced nature of information processing technology.

It is also crucial to effectively collect and process information, which includes two aspects of processing and transmission[4]. To establish a scientific and efficient information acquisition system, the new data information filtering and sorting in a timely manner, and to the original database and compatible effectively, provide more support for data search, at the same time as the information processing, more major is to use some of the data model, make the final evaluation and judgment for data value, So as to provide more support and help for the government and enterprise's information computing processing development ability. In the internet platform, the greater the amount of data information and the more complex the structure, the greater the challenge to the security of data information [5]. The development of science and technology has a two-sided effect. While providing people with a lot of convenience, hacker technology is also developing, and the number of virus software has also increased, which has brought troubles to people's lives to a certain extent. Technical personnel to improve information security is more in line with people's real interests, and can effectively avoid the occurrence of information security incidents. More and more professionals are committed to the research and development of information security technology, striving to build an effective security guarantee mechanism, maintain the security of network information, and effectively protect people's interests, which will contribute to social stability to a certain extent.

3.2. The development direction of computer information processing technology

With the continuous maturity of computer networks and the advancement of Internet communication technology, relying on the Internet platform for more data analysis has become a new development feature. In particular, people's demand for data usage has exceeded the specific requirements of computer hardware. Need, the hardware level of the computer itself cannot match the development of data. The emergence of cloud technology has effectively resolved this contradiction, especially as cloud computing itself. On the one hand, it can make the whole data transmission more convenient, and on the other hand, it can form a more powerful cloud computing network. This kind of network technology, will have more efficient processing ability, also can achieve timely feedback, which provides more help and support for the development and effective application of computer information processing technology, such a development trend will also become an important development trend of computer information processing technology.

The improvement of secure information technology will also be a very important development trend. In the era of big data, any data system can realize effective data sharing by means of the Internet.
This means that as long as access to the Internet, there is an opportunity to obtain a large amount of data, so comprehensively improving its security performance will become an important development trend in the future. On the one hand, the single data unit should be formed into a whole development system to effectively process and improve; On the other hand, the corresponding security software should be timely developed to provide effective support for the development of the whole computer system and the protection of Internet security have further improved the level of computer technology and promoted the development of data processing technology in the production process.

4. Conclusions

Developed technology makes the analysis technology of big data affect our life. In this information age, computer information processing technology combined with big data also presents a good future development trend. With the continuous popularization of Internet plus technology, cloud computing, cloud storage, network video communication and other technologies are also developing rapidly. In the future, cloud computing is likely to become the core of various computer information technologies. Cloud computing can help computers to greatly improve the speed of operation, so that computer information technology is more perfect. On the other hand, in the application process of cloud computing technology, data can be processed with characteristics, programming and avoidance in combination with different properties of different data, the benefits of large database can be reflected, and it is also helpful to computer technology. Moreover, the application of cloud technology can comprehensively expand the field of computer information technology, increase the storage of computers, maximize the efficiency of data processing, and constantly meet people's requirements for computer information processing.

Acknowledgments

The author is grateful to the Editors and the referees for their insightful comments that helped to significantly improve this article.

References

[1] White Tom. Hadoop: The Definitive Guide (1st ed) [M].USA: O'Reilly Media. Inc. 2009.
[2] Mu Q. Discussion on the application and practice of computer information processing technology under the background of big data [J]. Network Security Technology and Application, 2019 (11) :65-66.
[3] Data mining-PAM (K-Medoids) clustering algorithm learning_Pandora Galaxy column-CSDN blog https://blog.csdn.net/liuh845633242.
[4] Jia Conghui. Research on the penetration of computer information processing technology in the background of big data era [J]. Journal of propagation power, 2019,3 (28) :270.
[5] Gong Xiayi,Li Bohu,Chai Xudong,et al.Survey on big data platform technology [J]. Journal of System Simulation,2014,26(3):489-496.