Retraction

Retraction: Research on Optimization of Coding and Modulation Information Processing Technology in the Era of Big Data (*J. Phys.: Conf. Ser.* **1992** 022176)

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This article has been retracted by IOP Publishing following an allegation that raises concerns this article may have been created, manipulated, and/or sold by a commercial entity.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 Optimization of Coding and Modulation Information Processing Technology in the Era of Big Data

Ao Liu\(^{1,\ast}\), Chengliang Jiang\(^{1}\)

\(^{1}\)School of Telecommunications Engineering, Xidian University, China, 710071

\ast Corresponding author e-mail: jsu_1@stu.xidian.edu.cn

Abstract. With the continuous development of China's science and technology, the economic level is also constantly improving. In today's big data environment, the computer field has made a comprehensive breakthrough, and computer technology has been widely used in all walks of life. At present, the arrival of the information age has brought new opportunities and challenges to social development. Currently, people have many ways to obtain information. In the era of big data, various kinds of information are increasing, and all kinds of complex information will affect people's work and life. Therefore, there is an urgent need to optimize and innovate the coding and modulation information processing technology to improve the efficiency of information processing in the context of the big data era, so that people can efficiently obtain useful information from it to meet the needs of people's development. Actively explore the research of computer information processing technology, change the traditional information processing pattern, improve the accuracy of information needed in the era of big data, and promote the development of modern society and economy. This article mainly analyzes the current situation of the coding and modulation information processing technology under the background of the big data era, and discusses how to optimize and enhance the coding and modulation information processing technology under the background of the big data era. And try to find a new way out for the coding and modulation information processing technology in the current era based on the results of these studies. Experimental research shows that by controlling the maximum number of iterations, code length, code rate and other conditions of the encoding and decoding algorithm of LDPC codes, the coding efficiency can be effectively improved, and the coding and modulation information processing technology can be optimized.

Keywords: Big Data, Coding and Modulation, Information Processing, Optimization Research

1. Introduction
With the development of the times and the continuous progress of information technology, people's production and lifestyle have undergone tremendous changes. Under the background of the big data era, the information and data generated by people in production and life will be acquired and recorded,
which greatly meets the needs of business applications and life [1]. Currently, coding and modulation information processing technology is facing new challenges. How to deal with valuable information in big data more effectively increases the difficulty of information processing. The whole society is undergoing development and reform [2]. Traditional coding and modulation information processing technology has been difficult to meet the needs of a large amount of information and data processing, so it is necessary to improve, innovate and develop research [3].

In the current era of big data, the application range of computers is becoming wider and wider, and the demand for computer information processing technology is also higher and higher. People are also a big challenge [4]. The biggest feature of big data is that it has massive data, and massive data is the key to the information storage problem. The society is constantly developing, and all walks of life are also developing. As a result, data is updated all the time, and the types of resources are more abundant, so it is very difficult to store data and information. Traditional computer information processing technology can only store a small number of data information [5]. In the context of big data, traditional information storage technology and space can no longer meet the needs of massive data storage. If you continue to use traditional information coding technology to store data, it will affect the progress and development of society [6]. Therefore, it is necessary to in-depth research, develop new information coding technology, effectively reduce the storage space, and update it in time with the update of data information, filter effective information, and improve the efficiency of information utilization. In the process of development of computer information processing technology, there are still many problems, which hinder our country's informatization process. Therefore, relevant personnel must take certain countermeasures, and in-depth study and discussion of computer information processing technology, and then continue to promote the development of information technology in our country, making our country a leading country in science and technology [7].

Information is neither matter nor energy. This is the general term for the content that humans exchange with the external environment when they adapt and perceive the external environment. You can think of information as a signal for people to communicate with the outside world [8]. In the current global development model, human thinking has reached a very high horizon. The development of everything has certain and inviolable objective laws. Through computer information processing technology, information is collected from various channels, and the information is processed and analyzed through coding technology, which can greatly facilitate the storage and calling of information [9]. It is worth noting that in the context of the era of big data, information can be called "all-inclusive". Any trivial matter may become the target of analysis. It can grasp the key points, explore the objective laws behind the success of others, and then target them. To a certain extent, the distance between the two parties can be shortened quickly, thus making our own success faster. However, computer information processing technology based on the background of big data is also a "double-edged sword". In the actual application process, we must pay attention to the grasp of the scale and carry out it reasonably within the legal framework. There are still many difficulties to be overcome for information processing technology in the future [10].

2. Method

2.1. Build a coding system

If a function $g(f)$ satisfies the following relationship:

$$
\int_{-\infty}^{\infty} |g(t)|^2 \, dt < \infty \quad (1)
$$

Then the space formed by the function $g(t)$ is called the square integrable space, that is, $g(t) \in (R)$. From this formula, we can see that $g(t)$ is an energy finite function.

If the function $y(f) \in (t)$, its corresponding Fourier transform is, when the condition is met:

$$
C_{\psi} = \int_{-\infty}^{\infty} \left| \hat{\varphi}(\omega) \right|^2 \, d\omega < \infty \quad (2)
$$

It is called a basic wavelet function. After the function is translated and stretched, the result is a
wavelet sequence:
\[ \varphi_{a,b}(t) = \frac{1}{\sqrt{a}} \sigma \left( \frac{t - b}{a} \right) \]

2.2. Reduce information coding errors
There are two types of errors involved in information coding. One type is intra-frame bit errors caused during wireless transmission, which are called physical layer errors here; the other type is transmission errors caused by layers above the physical layer, which are called upper layer errors. When considering joint optimization, a method is needed to measure the transmission error of the entire system, that is, a parameter is required to unify the physical layer error and the upper layer error for easy measurement and comparison. However, the physical layer measures the probability of error. The probability of bit error is BER. PER is more than a measure of the physical layer. It is necessary to use PER as the measurement index of the transmission error of the entire system, which requires the corresponding PER to be obtained from the BER, then:
\[ \text{PER} = 1 - (1 - \text{BER}) \]

Np is the number of bits contained in a packet, but in fact, the error probability that may be experienced between bits in the same packet is generally different. In addition, if channel coding is used, the error probability between bits will be correlated, and the hypothesis will be invalid. In the following, for the two cases with channel coding and without channel coding, the estimation formula of error transmission probability is roughly calculated. The calculation method of QAM modulation PER under AWGN channel has been given. Using this method, first find out The error probability calculation formula of the transmission mode using error correction coding. Then in the I-order QAM modulation, the error probability of the k-th symbol is:
\[ \text{PER}_{\text{QAM}} = 1 - \left( \frac{1}{\log_2 i} \right) \left( 1 - P(x) \right)^{N_p} \left( 1 - \frac{1}{\log_2 i} \right)^{\frac{N_p}{j}} \]

It can be obtained uniformly with the given PER expression. According to the calculation formula of PER, the calculation formula of PER in each mode can be obtained:
\[ \text{PER}_i(y) = \begin{cases} 1, & 0 < y \leq y_n \\ a_n \exp(-g_n y), & y \geq y_n \end{cases} \]

2.3. Cloud computing information processing
In computer information processing, we need to pay attention to the content of the cloud computing field. This is to ensure the necessary content of the data processing content of the information system. With the strengthening of data processing, cloud computing and big data have gradually developed into mutual roughness. The relationship between the two can complement each other, which greatly reduces the need for computer support for computing power in the environment of the big data era, which may lead to the disadvantages of high requirements for computer hardware performance. Cloud computing can be used. The platform decomposes the system data in a hierarchical manner, and can complete the collected analysis in the internal server. For this reason, for the user terminal, we are actually the recipient of the final information, so not only can the user have a better experience can also reduce the cost of users on computer configuration.

3. Model establishment

3.1. LDPC information coding model
The LDPC code uses an iterative probability code, which is a bit-by-bit soft decision method.

Given a N-length LDPC code, the codeword X= represents a group of N-length information nodes, and Z= represents a group of M-length check nodes. The information node is called the parent node of all the inspection nodes connected to it, and the inspection node is called the child node of all the
information nodes connected to it.

Define the set of receiving vectors as: $S$ is defined as the conditional probability of the information bit in the received bit and $S$ in the event that the information node $X$ satisfies all the included test equations is 0 or 1. Consider the following relationship:

$$\frac{P_S(x_i = 0 \mid r, S)}{P_S(x_i = 1 \mid r, S)} \geq 1$$  \hspace{1cm} (7)$$

When the above formula is true, the received bit value $= 0$, otherwise $= 1$. We will introduce LDPC decoding in detail below. Define the probability of an even number of 1s in an $n$-length binary sequence as follows:

$$\frac{1}{2} \left(1 + \left(1 - 2P_j^i\right)\right)$$  \hspace{1cm} (8)$$

Where the $P_j^i$ probability that the $j$th bit in the sequence is 1. Defined as a $R_k^j$ test message, which means the probability of satisfying the $j$-th test equation when $= a$. When $x_j = 0$, we can get

$$R_k^j = \frac{1}{2} \left(1 + \left(1 - 2P_j^i\right)\right)$$  \hspace{1cm} (9)$$

Corresponding $x_j = 1$ to $= 1$, the above probability is

$$R_k^j = 1 - R_k^j$$  \hspace{1cm} (10)$$

The left side of the inequality can have the following forms:

$$\frac{P_r(x_i = 0 \mid r, S)}{P_r(x_i = 1 \mid r, S)} = R_k^i \hspace{1cm} (11)$$

The left side of the inequality can have the following forms:

$$\frac{P_r(x_i = 0 \mid r, S)}{P_r(x_i = 1 \mid r, S)} = R_k^i \hspace{1cm} (12)$$

In this way, in the BP algorithm, in the second iteration of the distribution probability of the initially obtained information message, each information node passes the current to its corresponding child node $r$, through the received, update, value, and then pass to its corresponding parent node to update Value; After the iteration, the information node judges through all the received values, and finally gets the decoded output. Assuming that the information passes through the AWGN channel after BPSK modulation, the noise power spectral density is $N/2 = \sigma_n$, the maximum likelihood probability function of the received bit $r$, is expressed as follows:

$$P(r \mid x_i = 0) = \frac{1}{\sqrt{\pi N_0}} e^{\frac{1}{\sqrt{\pi N_0}}}$$  \hspace{1cm} (14)$$

From the Bayesian formula:

$$Q_j^0(t + 1) = K_j \left(1 - P_j^i\right)R_k^i(t + 1)$$  \hspace{1cm} (16)$$

If the check matrix $H$ is known, the code word obtained by the judgment is described as $x$, and the result of calculating $T$ is all zeros. If it is, the decoding is correct; otherwise, it is judged whether $t$ reaches the maximum value. If $T$ reaches the maximum value, it indicates the setting The number of iterations is completed, and the codeword is output as the decoding result; if there is no $t+1$, then it is transferred to continue iteration. At this point, the entire decoding process is over. The sum product decoding algorithm can be obtained after the likelihood ratio measurement is introduced to the BP algorithm, and the minimum sum algorithm can be obtained after further improvement.
algorithm has the same decoding principle, but the calculation process is more simplified.

4. Evaluation results

4.1. Analysis of experimental research survey results
Since the maximum number of iterations for decoding is 30, the number of iterations set in this experiment is also within the range of values, and the number of iterations is set to 2, 4, 6, 8, and 10. 12 times, 14 times, 16 times when the program consumes time with the same number of threads. And tested the time consumed by the program when decoding at a bit rate of 0.125, a bit rate of 0.25 and a bit rate of 0.5 in different thread numbers.

| Number of iterations | 2   | 4   | 6   | 8   |
|----------------------|-----|-----|-----|-----|
| Time                 | 0.674| 0.624| 0.582| 0.547|
| Number of iterations | 10  | 12  | 14  | 16  |
| Time                 | 0.538| 0.524| 0.521| 0.516|

From the data in the table, it can be known that as the number of iterations increases, the time consumed by the program is gradually decreasing, but when the number of iterations reaches a certain level, this downward trend is also slowing down and gradually calms down. It can be seen from the data that the number of iterations of decoding needs to be controlled to a certain extent. Choosing an appropriate number of iterations can reduce the running time of the program and speed up the decoding. Combining the data in the table and providing a priori information through deep learning information processing, an iterative decoding method is proposed, and the information coding model is improved, and a new information coding data model based on the LDPC algorithm is designed.

Figure 1. The performance of three different bit rates.

Figure 2. Code length improves the performance of the compilation algorithm.
As shown in Figure 1, the experiment has made a performance comparison under three different code rates. The results are shown in the figure. The three code rates of 0.125, 0.25 and 0.5 are respectively analyzed by experiments, 0.125, 0.25 and 0.5. The code uses 22, 27, and 19 features respectively. The data in Figure 1 shows the performance of the three code rates under the same conditions. It can be seen in the figure that the 0.125 code rate is the lowest performance of the three cases, and the highest performance is the 0.5 code rate. Time efficiency is an important performance of the LDPC processing algorithm. At the same time, the time efficiency of the algorithm in the training phase and the test phase is considered. Figure 1 shows the experimental results of the three code rates and test time. In general, the 0.125 bit rate of the linear core achieves the lowest computing time. It can be known from the data analysis in Figure 2 that the code length also has an impact on the performance of the compilation algorithm. The longer the code length, the more complete the information expressed, and the higher the efficiency of the compilation algorithm. Therefore, in information processing, a certain degree of lengthening of the code length is also beneficial to coding. The optimization of information coding needs to speed up the time of information processing. While reducing the time, it is also necessary to improve the accuracy of the information. A reasonable use of the code rate of the code can greatly improve the efficiency of coding and speed up the information transmission.

5. Conclusion
Aiming at the relatively high computational complexity of traditional BP decoding algorithms, and combining other decoding algorithms, this paper studies the encoding and decoding algorithms of LDPC codes. The LDPC code is researched and analyzed, and the decoding algorithm is discussed in terms of the maximum number of iterations, coding length and coding rate. The results show that within a certain range, the greater the number of iterations, the lower the bit error rate, but when the number of iterations reaches a certain level, the increase in the number of iterations does not significantly increase the improvement effect of the compilation algorithm. Under the same conditions, the longer the code length, the performance of the compilation algorithm can be improved to a certain extent. The greater the bit rate, the greater the bit error rate. We can optimize the coding and modulation information processing technology by increasing the maximum number of iterations, code length, and reducing the code rate.

References
[1] Wang Y, Kung L A, Byrd T A. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations[J]. Technological Forecasting & Social Change, 2018, 126(JAN.):3-13.
[2] Wang X, Zhang Y, Leung V C M, et al. D2D Big Data: Content Deliveries over Wireless Device-to-Device Sharing in Large Scale Mobile Networks[J]. IEEE Wireless Communications, 2018, 25(1):32-38.
[3] Yudong C, Yuejie C. Harnessing Structures in Big Data via Guaranteed Low-Rank Matrix Estimation[J]. IEEE Signal Processing Magazine, 2018, 35(4):14-31.
[4] Gu K, Tao D, Guo J F, et al. Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data[J]. IEEE Trans Neural Netw Learn Syst, 2018, 29(4):1301-1313.
[5] Vernon P A. Speed of information processing and general intelligence[J]. Intelligence, 2018, 71(1):33-70.
[6] Watson H, Guevara M, Vilstrup H, et al. Improvement of hyponatremia in cirrhosis is associated with improved complex information processing[J]. Journal of Gastroenterology and Hepatology, 2019, 34(11):1999-2003.
[7] Lyu X. Optimization Research on the Voltage of Distribution Grid with HHPV Accessed Based on GA[J]. International Core Journal of Engineering, 2020, 6(4):138-146.
[8] Wang L, Guo S, Wei Y, et al. Optimization research on the lubrication characteristics for friction pairs surface of journal bearings with micro texture[J]. Meccanica, 2019,
[9] Chen G, Li J. A Research Mode Based Evolutionary Algorithm for Many-Objective Optimization[J]. Chinese Journal of Electronics, 2019, 28(4):764-772.

[10] Wang L, Guo S, Wei Y, et al. Optimization research on the lubrication characteristics for friction pairs surface of journal bearings with micro texture[J]. Meccanica, 2019, 54(8):1135-1148.

54(8):1135-1148.