Based on the Software Complexity Measurement of Complex Networks under Big Data Technology

Xiaobin Hong1,*

1Guangzhou College of Technology and Business, 510000, Guangzhou, Guangdong, China

*Corresponding author e-mail: xiaobin12004@gzgs.edu.cn

Abstract. With the development of the times, computer technology is booming, so the network is becoming more and more complex, software design is becoming more and more complex, because of the protection against a variety of internal or external risks. The internal risk is that the traffic carried by the system is too large to cause the system to crash or the system to crash caused by the code operation error, and the external threat is that hackers use computer technology to break into the system according to security vulnerabilities, so the purpose of this paper is based on big data technology, the software complexity of complex networks is measured and studied. With the consent of the school, we used the school's internal network data, and after consulting the literature on the complex construction and analysis of complex networks and software, modeled and analyzed it using the improved particle group algorithm. The experimental results show that there is a certain correlation between complex network and software complexity. Because complex networks determine that software requires complex construction to withstand potential risks to keep the software running properly.

Keywords: Big Data Technology, Particle Group Algorithms, Software Complexity Research, Complex Network Research

1. Introduction

In this rapidly changing society, new technologies are discovered and new ideas are put forward all the time[1]. So, we should keep pace with the times. Because of the development of the times, the network is becoming more and more complex, and all kinds of data are full of it. Only with the huge data flow can break down the firewall of most enterprises [2]. And some highly skilled hackers can do whatever they want in the network world, so we conduct a comprehensive study on the complexity of
complex networks and software based on big data technology to explore the relationship between them and make a good layout [3].

Big data technology is not only a technology, but also a symbol of an era [4]. It shows that our era is formed by the alternation of various data streams. Because of the advent of the information age, all people are communicating with others and exchanging their ideas through intelligent devices through the network, so there are huge data streams running alternately every day [5]. But too much data flow is also easy to cause many serious consequences, because too large data flow, it is easy to store some error codes or a Trojan horse program, to invade other people's systems, to achieve some unknown purposes [6]. So, in this era, the construction of firewall is very necessary, and is the top priority of security system. Because now all the information is stored in the internal database. If something goes wrong, it will leak not only the user's data, but also the company's secrets, which will make a company lose its vitality [7-8].

So, in order to better protect data security, we mainly analyze the complexity of the software by analyzing the complexity of the network, and then relatively observe their situation, whether there is relevance [9]. Then, through experiments, the data complexity, data flow density and software system complexity are increased respectively, and their relationship is comprehensively analyzed to obtain the experimental results [10].

2. Improved particle group algorithm

Take bird foraging, for example, assuming that in an N-dimensional space environment, there are m particles, and the optimal foraging position passing through the particle i calendar can be expressed as \( P_{\text{pi}1}, P_{\text{pi}2}, \ldots, P_{\text{pid}} \), and the optimal historical foraging position found by the particle group can be represented. The speed and position change formula for the d \( i \)-dimensional mass \((1 \leq d \leq N)\) of the i particle in the target search space is:

\[
v_{id}(t + 1) = wv_{id}(t) + c_1r_1(p_{id}(t) - x_{id}(t)) + c_2r_2(p_{gd}(t) - x_{id}(t))
\]

\[
x_{id}(t + 1) = v_{id}(t + 1) + x_{id}(t)
\]

\[
w = w_o - (w_s - w_0)\frac{g}{g_{\text{max}}}
\]

In the series, \( v_{id} \) is the dD-dimensional part of the particle i-velocity vector, \( x_{id} \) is the d-dimensional part of the particle i-position vector, \( c_1 \) and \( c_2 \) are the learning factors, \( r_1 \) and \( r_2 \) are random numbers within the 0,1, \( s \) is the maximum weight coefficient, \( e \) is the minimum weight coefficient, \( G_{\text{max}} \) is the largest evolutionary algegae, and \( g \) is the current evolutionary algeausal.

In the particle group algorithm, the information obtained by each particle is determined by the historical optimal value of the particle individual and the global optimal solution of the whole particle group, and as the particle group evolves, the particle vitality decreases, the particle velocity becomes 0 and it is easy to get into the problem of local extreme values. In order to strengthen the range of particle search and give diversity to the speed of particles, the most obvious particle information, i.e., information that conforms to the largest particle in the particle group, is introduced. The introduction of active factors enhances the vitality of the current velocity component. The update formula for particle \( |p_{\text{besti}} - s_{\text{besti}}| \) velocity is:
\[ v_{id}(t + 1) = w v_{id}(t) + c_1 r_1 (p_{id}(t) - x_{id}(t)) + c_2 r_2 (p_{gd}(t) - x_{id}(t)) + c_3 r_3 (s_{id}(t) - x_{id}(t)) \]  \( (4) \)

Sid is the active factor introduced, pbesti is the best solution in current history, and spbesti is the best solution in the history of the last generation.

3. **Experiment build**

3.1. *The selection processing of the experiment*

In order to do the experiment better, we chose a computer room to do the experiment, and have been collecting data, and then calculating and coupling them, modeling the data we need.

3.2. *Implementation of the experiment*

So, we borrowed some data from the school, and then we made a simple system to deal with the data flow, and then gradually increased the data flow, slowly to observe the situation of the software. Then improve the complexity of the system, slowly improve each other, and then a little analysis, and finally get the experimental results.

4. **Evaluation results**

4.1. *Experimental results*

**Table 1.** compares the complexity of the data flow to the complexity of the software

|                | The complexity of the data flow | Software complexity |
|----------------|---------------------------------|---------------------|
| Phase 1        | 46%                             | 51%                 |
| Phase 2        | 57%                             | 60%                 |
| Stage 3        | 81%                             | 82%                 |
| Stage 4        | 96%                             | 98%                 |

**Figure 1.** About the comparison between the complexity of data flow and the complexity of software
Based on the data results from Table 1, we found that: An experiment we divide into four phases, then step by step to increase the complexity of the data flow and software complexity, and then compare them to this possible serious vulnerability, general vulnerability, and security situation. It is concluded that as the complexity of the data increases, so does the complexity of the software, as well as its security. Figure 1, Figure 2, is a diagram based on the results of the experiment to show a more intuitive image.

4.2. Complex networks

Complex networks are defined by Mr. Qian Xuesen. Networks with self-organizing, self-similarity, attractors, small worlds, and no scales of some or all nature are called complex networks. At the macro level, a complex network is a particularly complex network with several characteristics. First, there are many nodes and there are many characteristics of the network structure. Second, web pages and links may be disconnected or appear at any time, the network structure has been changing, is dynamic. Third, there is diversity in the connection. There may be directionality. Fou, You can't think linearly, it's moving in a nonlinear way. Five, any node inside can represent any different thing or organization. Sixth, complexity can be superimposed to the point where it is more complex.

Complex networks typically have three characteristics. The first feature is the small world. It shows that although there are many nodes in a complex network. However, there is a shortest path between each node, that is, the shortest straight line between two points. The second feature is clustering. It means that the network is the same as a group, slowly clustered, grouped. The more advanced the cluster, the more connected the network nodes are. The third feature is the power ratio. Refers to the concentration or closeness of the cluster.

4.3. Software complexity

Software complexity follows the principle that simplicity is reliable, and complexity is constantly changing. In the 1970s, software systems became incredibly complex. Both development and maintenance become extremely complex and difficult, and the costs are particularly high. So, then people came up with a way to modularize the software for development, testing, and maintenance. There are three kinds of software complexity. They are modules, classes, and programs. Because the composition of software is complex and diverse, there is a particularly large amount of code, we have
to find one or several lines of code from these lines of unreasonable failure to run is very difficult, which requires a lot of human and material resources. So, then people made up the software into blocks. It is then developed, maintained, and processed separately. Just like the composition of the software we see now.

4.4. Software metrics

Its purpose is to improve the process of software development, develop high-quality products, and promote the success of the overall project. Software metrics are currently divided into two factions, one that believes software can be measured and the other that software cannot pass measurement analysis. But the current research mainstream is considered measurable, so we think that in the current situation, the complexity of the software is not very high, can also be measured and analyzed. If we can, we want the software to be measured and analyzed all the time. Because measurement analysis is equivalent to a software composition of a proposal, we can use real-time measurement analysis to control the software in real time, which can help us make better software for service to success.

4.5. Big data

The most important thing about big data is that it transforms information. And big data technology must have cloud computing, because the amount of big data is too large, a single computer cannot process, can only use distributed computing, the use of cloud computing to comprehensively process data, its classification, and then statistics the required data, it is stored and converted into the information we need. With the advent of the era of cloud computing, the value of big data is getting higher and higher, has become a hot thing for enterprises.

Broadly speaking, cloud computing is a service that leverages computer technology and the Internet, and we call one of their pools of shared resources the cloud. Cloud computing brings together many data computing resources, and automated management through the environment allows resources to be provided quickly. Cloud computing is not a new technology, but a new concept. At its core is the Internet, which provides fast and secure storage services. At present, cloud computing has become a new era - the era of the cloud. And cloud computing has become a new revolution in computers since it was proposed a decade ago. And because of his appearance led to a new change in society as a whole. Cloud computing is characterized by its virtualization and does not require physical control. This is his greatest advantage. Then there's the flexibility, reliability, and price/performance ratio (as opposed to physical storage), and the ability to extract anywhere, anytime, which is an advantage. But it also has the disadvantage that the data stored in it may also be compromised and leaked. And because of the storage of cloud computing, people can use other people's personal information for information theft. And because of the huge data resources of cloud computing, it may appear that some people bring viruses to cloud computing, causing the entire cloud computing system to crash and collapse completely. So, we need to be careful about using cloud computing.

Big data also has a variety of structures, namely structured, semi-structured and unstructured data. Today, most of the data in the enterprise is unstructured. With the development of the times, big data will gradually create a more convenient life and more wealth for mankind. Now this society is a high-speed development of society, science and technology has been developing. The exchange of
information is getting faster and faster, life is more and more convenient, and big data will be available in the future. Including now such as Taobao, JD.com and other types of e-commerce, as well as the United States group, hungry and other types of takeaway services are using big data technology to mine user data, divided according to people's preferences to provide people with what they want. And in the future the trend for big data should be to become a very valuable resource and then further combine with cloud computing. Then, with the breakthrough of theory, we can create more new big data-based technologies to better use big data to analyze the nature of processing things. And in the future, a special discipline may be set up for this purpose. However, due to the proliferation of data, the possibility of data leakage in the future will be greatly increased, so the most important thing in the future is to protect their own information privacy. Only by protecting the privacy of your information can you not be afraid to start something else. And in the future, it is bound to be an ecosystem of the data age and the Internet of Things to improve our modern life.

5. Conclusion

With the development of the times, the application of complex networks is becoming more and more extensive. To keep our data safe, we often set up multiple nodes on the network to confuse others and mask real data. Although the number of complex network nodes is large, but we have a set of applications belonging to their own, can quickly and easily find the best nodes to perform operations. Therefore, according to the experiment, the complex network and software complexity are necessarily linked, the more complex the software, the network is also more complex.

Acknowledgments

Project: The characteristic Innovation Project Natural Science in Guangdong university-Research from Big Data on analysis of User's Portrait and Behavior (2019KTSCX258).

References

[1] Jain N K, Nangia U, Jain J. A Review of Particle Swarm Optimization[J]. Journal of the Institution of Engineers, 2018, 99(4):1-5.

[2] Li M, Chen L, Xu Y. Extracting core questions in community question answering based on particle swarm optimization[J]. Data Technologies and Applications, 2019, 53(4):456-483.

[3] Saha, Sriparna, Das, et al. Exploring differential evolution and particle swarm optimization to develop some symmetry-based automatic clustering techniques: application to gene clustering[J]. Neural computing & applications, 2018, 30(3):735-757.

[4] A Z W, B H W A, B S W A, et al. Linear mixed-effects model for longitudinal complex data with diversified characteristics[J]. Journal of Management Science and Engineering, 2020, 5(2):105-124.

[5] Templ M, Gussenbauer J, Filzmoser P. Evaluation of robust outlier detection methods for zero-inflated complex data[J]. Journal of Applied Statistics, 2020, 47(7):1144-1167.

[6] Yajie L, Qicheng C, Sundhararajan M, et al. Complex data study on mechanical fault
diagnosing[J]. Journal of Intelligent & Fuzzy Systems, 2018, 34(2):1169-1176.

[7] Wang D, Tan D, Liu L. Particle swarm optimization algorithm: an overview[J]. Soft Computing, 2018, 22(2):387-408.

[8] Wang Z, Cai J. The path-planning in radioactive environment of nuclear facilities using an improved particle swarm optimization algorithm[J]. Nuclear Engineering and Design, 2018, 326(JAN.):79-86.

[9] Qamar N, Malik A A. Impact of Design Patterns on Software Complexity and Size[J]. Mehran University Research Journal of Engineering and Technology, 2020, 39(2):342-352.

[10] Akundi A, Smith E, Tseng T L. Information entropy applied to software-based control flow graphs[J]. International Journal of System Assurance Engineering and Management, 2018, 9(5):1080-1091.