

Safe retrieval of sensitive information on mobile internet under multi-stage attack

Xiang Zhang

Anshun Internet Public Opinion Monitoring Center, Anshun 561000, China

Keywords: multi-stage attack; sensitive information; secure retrieval

Abstract: The data era has emerged with the quiet landing of the cloud computing and mobile Internet era, and gradually moved from "why to use big data technology" to "how to use the current big data core technology to comprehensively strengthen future network security strategies and tactics". Sensitive information retrieval is a hotspot in secure multi-party computing research, and the existing sensitive information retrieval methods cannot effectively protect the user's query information. Semantic retrieval based on ontology is one of the effective ways to improve the retrieval accuracy of information systems. Concept cloud is the application of cloud model in the information field, which is the arrangement and aggregation of concepts in the ontology in the page according to a certain algorithm. When designing some information systems, there are often users who require that the data in the database can be queried in any combination, which is convenient and easy to use and does not require high database expertise. The development of modern information technologies such as communication technology, computer technology, network technology, and audio-visual technology has been integrated with each other, broadening the scope of information transmission and application. For an information system, how to meet the information needs of users is a very important goal.

1. Introduction

In recent years, with the rapid development of the Internet, the information on it has become more and more abundant, and people's requirements for resources are also increasing. Resource discovery on the Internet is a promising research direction. Users can discover systems from different resources. to obtain the required information [1]. With the emergence of a large number of sensitive information and the multi-party collaborative computing... Cloud computing provides a platform for big data processing, storage and computing, and improves the processing speed of big data. To protect data privacy, sensitive information needs to be encrypted before uploading to cloud servers [2]. However, if the user wants to retrieve information from the server, it needs to be decrypted before retrieval, which will lead to the leakage of the plaintext information to the untrusted cloud server, which will damage the security of the information. The sensitive information retrieval problem has broad application prospects in the field of secure multi-party computing, and plays a very important role in many practical applications [3]. The cloud platform will modify, delete or destroy data to a large extent. People pay more and more attention to the security of sensitive information in the network. It is necessary to study the secure retrieval of sensitive information in wireless local area networks. Provide a basis for ensuring the security of information in the wireless local area network. The key to secure information retrieval in encrypted domain is to keep the distance between information features before and after encryption unchanged. Third parties can be brought in [4]. Since both the querying parties perform hash transformation on the input before submitting the data, the third party at a loss cannot obtain any information secretly input by the querying parties, and thus does not know the final result. At the same time, only one database copy is generated during the query process, which effectively protects the security of sensitive information.

2. Information Security Retrieval

2.1. Basic Definition of Information Security Retrieval

WWW (WORLD WIDE WEB) has been developing at an alarming rate since its inception, the amount of information is increasing, more and more people visit it, and people use it as a tool for study, work, entertainment, etc. [5]. Under the traditional keyword-based information retrieval method, due to the lack of unified semantic description of information resources, it is difficult for users to find resources related to their needs. The main reason is that the semantic sharing of information resources cannot be achieved [6]. The fundamental solution to this problem is to improve information retrieval from the search based on keywords to the level based on semantics. WWW intelligent retrieval system consists of four parts, UI AGENT, BROKER, SPIDER AGENT and LEARNER. The overall structure of each part and the relationship between them are shown in Figure 1 below.

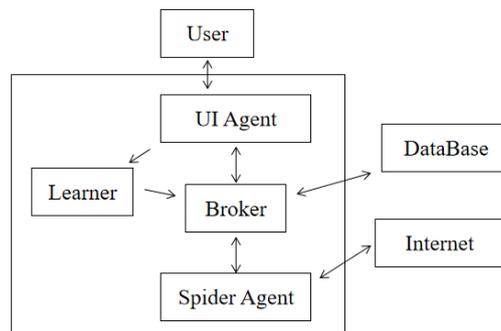


Figure 1 The structure of WWW intelligent retrieval system

Semantic retrieval based on ontology is one of the effective ways to improve the retrieval accuracy of information systems. Concept cloud is the application of cloud model in the information field, which is the arrangement and aggregation of concepts in the ontology in the page according to a certain algorithm. Because the development of Internet information often cannot keep up with the growth of Internet information, search engines can find documents related to keywords for people, but the returned results are often too many documents and low hit rate [7]. When people conduct information retrieval, they always expect to find all documents related to the topic of interest, that is, they hope to obtain a higher retrieval rate, and they do not want to include documents that are not related to the topic, that is, they expect to obtain higher retrieval accuracy. . It is very difficult to achieve both at the same time. The World Wide Web provides users with an open information sharing resource platform. Compared with the traditional closed information system, people can publish and share information resources globally at the fastest speed [8]. High-performance information retrieval based on Web mining is proposed to address this problem. The goal of the research is to fully collect the personalized information of Lichuan Chuanhu, collect Web information through flexible means such as Ganghu interest guidance or rich query modes, and fully Lichuan network information, so as to improve the accuracy of Qibo, improve the retrieval quality, and meet the , } The specific query requirements of J households. With the rapid growth of the amount of information, it has also caused many problems such as "information loss" and "information overload". People find that it is not easy to find valuable information in this ever-changing and mixed information ocean. [9]. Because the premise of high-performance personalized information retrieval is to mine the access characteristics of users and classify users, the user's access behavior is stored in the web log, and the web log data needs to be preprocessed to be able to mine the interests of the state households. .

2.2. Principles of Information Security Retrieval

Computer information retrieval refers to the use of computers to store and retrieve information. Specifically, it means that people retrieve the required information from the database of the

computer retrieval system using specific retrieval instructions, retrieval words and retrieval strategies on the computer or the terminal of the computer retrieval network, and then display or display the required information by the terminal device. The process of printing [10]. The data in the database should be described completely and reliably. Now, the data dictionary method is generally adopted, which is also called metadata. The so-called metadata refers to the data used to describe and describe other data. Information retrieval mainly studies the representation, storage, organization and access of information. That is, according to the user's query requirements, the relevant information materials are retrieved from the information database. Now, the Internet has been well utilized as a high-tech information release tool. Robust operating systems and literature databases provide users with great convenience [11]. The collection, arrangement, accumulation, filing and updating of data information make the literature database have extremely high scientific value. The format of metadata is generally a data description method determined by authoritative organizations after users study the nature of data and follow data specifications. Through these descriptions, users can perform some basic retrievals on the data. Information retrieval has developed from manual keyword indexing to computer automatic indexing of full-text information retrieval, automatic information summarization, and automatic information classification, and is developing in the direction of natural language processing [12]. In the field of information retrieval, English information retrieval develops rapidly. In order to realize computer information retrieval, a large amount of original information must be processed and stored in the computer in the form of a database in advance, so computer information retrieval in a broad sense includes two aspects of information storage and retrieval. There are tens of thousands of technical documents on the site of automation professional technical documents, and the work of the frequently updated database includes the content of adding thousands of documents each time [13]. Therefore, the user's efficient query of the database becomes an important issue. Since data is often dynamic and will change at any time, if there is no such mechanism, users may not be able to retrieve the data in the database through the data dictionary. The solution is to use a dynamic data dictionary, which can change with the user's needs.

3. Safe retrieval of sensitive information on mobile internet under multi-stage attack

3.1. Main functions of big data retrieval and security analysis

Network security is the core of ensuring the healthy operation of the national information society. Information security service and technology providers are required to be good at using advanced IT technology to predict, identify, and resolve network security risks. Identify the main factors and information security issues that may affect business development in the future from complex data, so as to further provide technical support for the security supervision of information security level protection. The development of information technology has greatly improved the dissemination speed and transmission efficiency of information and data. Although it has promoted the prosperity and development of the network world, it has also given some criminals an opportunity. The development and application of big data technology has had a profound impact on the governance model of the country, the decision-making structure of the enterprise, the business strategy of the business and the way of life of the individual. However, the accumulation of large amounts of data not only increases the risk of user privacy leakage, but also attracts more potential attackers due to the huge information and potential value contained in big data. The construction of a complete power information system can better provide favorable conditions for residential electricity consumption and enterprise services, and it is also helpful to further promote the informatization development of the power industry. Be sure to increase its protective measures to eliminate the security problems that arise. In the big data environment, security managers need to understand the related technologies of cloud computing and big data and the entire life cycle of retrieving and analyzing data, as well as the business processes, data processes and information systems of the security management organization or enterprise. The development of network information technology is a double-edged sword. Although users' use of new information technology has greatly

improved the speed of information dissemination and exchange, it has also increased the pressure of network security analysis work. At present, there are many users of the power information network security analysis platform. Therefore, a sound information security organization should be used as a guarantee to effectively control various aspects such as login access and transaction exit, so as to avoid security problems during use. Big data platform security mainly addresses identity authentication, data isolation, data encryption storage, big data platform boundary protection and auditing between big data components. The main key technologies are identity authentication, access control, data encryption and auditing. At present, there is still a lack of a complete big data security standard system in the world, and there is a lack of standard norms and guidance in privacy protection, data sharing, and data cross-border transmission.

3.2. Big Data Retrieval and Analysis Platform Architecture

Driven by the rapid development of information and network technology, people's perception, calculation, simulation, simulation, dissemination and other activities generate a large amount of data, which stimulates the vigorous development of cloud computing technology. With the help of cloud computing technology, people can successfully realize the retrieval and expansion of big data network information, and enhance the effect of Internet application. Online commerce, social networking, automated sensing systems, the use of mobile devices, and the use of scientific instruments generate vast amounts of data. Among them, the application of social networks (audio, video and pictures, etc.) and e-commerce make the data grow extremely rapidly. The correct processing of big data information by cloud computing and obtaining favorable demand information will have a huge impact on today's society. For this reason, the collection and retrieval of massive data by cloud computing has become a problem that people are working on. Pipe every field application department needs to be tailored, especially in the user domain and data collection, but the core technologies of personalization are similar, as shown in Figure 2.

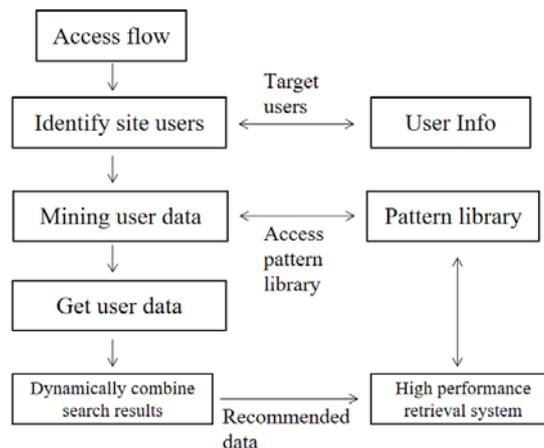


Figure 2 The workflow of the domain-oriented high-performance retrieval system

Big data is generated with the supercomputing power of cloud computing. The development of a large-scale distributed network application architecture model for big data and an efficient development and deployment method for application software is one of the main challenges facing the current big data management and application. Information retrieval service is an important function of the Internet information system. Due to the huge amount of information carried by the Internet, the operation of the information retrieval function is faced with a huge workload of data processing. Cloud computing technology, as an emerging big data technology, which can effectively support the operation of network information retrieval and expansion functions in the big data environment, and enhance the performance of the information retrieval system. Cloud computing is the development of distributed computing, parallel computing, and grid computing, and cloud is a resource pool containing a large number of available virtual resources. With the advent of the DT era, cloud computing technology has developed rapidly as a new technology trend, fundamentally changing the current way of people's work. Big data refers to: a large amount of

structured and semi-structured data. It has four characteristics. First, the volume is large, the second is the diversification of data types, the third is the low value density, and the fourth is the ability to process quickly. These four characteristics of big data indicate that they are stored, collected and retrieved differently from traditional ways. At present, the results of specializing in large-scale distributed application architecture for big data and efficient development and deployment methods for big data-oriented application software are rare. Research basis for large-scale distributed application architecture for big data.

4. Conclusions

In recent years, cloud computing has achieved rapid development due to its good elasticity, strong scalability, economical and practical characteristics, etc. In the field of big data research, there have also been some technical systems that apply cloud computing to big data. However, the serious lack of theoretical models limits the popularization and application of these technical systems. The big data network information retrieval technology based on cloud computing still needs to be improved and optimized. With the advent of the big data era, artificial intelligence technology can also be applied to network information retrieval, so that the retrieval task can be completed more intelligently. The massive, diverse and high-speed characteristics of big data have brought severe challenges to traditional technologies in data organization and management, data retrieval and query, data application processing, and application software development for small sample data processing. In the future development of information retrieval technology, it is necessary to continue in-depth research, optimization and innovation to obtain more intelligent and diversified retrieval results. The research of big data retrieval technology still has a long way to go.

References

- [1] Na S H, Kim K. Verbosity Normalized Pseudo-Relevance Feedback in Information Retrieval[J]. *Information Processing & Management*, 2017, 54(2):219-239.
- [2] Krallinger M, Rabal O, Louren?O A, et al. Information Retrieval and Text Mining Technologies for Chemistry[J]. *Chemical Reviews*, 2017, 117(12).
- [3] Mayr P, Frommholz I, Cabanac G. Bibliometric-enhanced Information Retrieval[J]. *Scientometrics*, 2018:1-1.
- [4] Ma C, Wei X, Chen F, et al. A Content-Based Remote Sensing Image Change Information Retrieval Model[J]. *ISPRS International Journal of Geo-Information*, 2017, 6(10):310-310.
- [5] Kim H, Cha M, Kim B C, et al. Part library-based information retrieval and inspection framework to support part maintenance using 3D printing technology[J]. *Rapid Prototyping Journal*, 2019, 25(3):630-644.
- [6] Abdulahhad K, Chevallet J P, Berrut C. Logics, Lattices and Probability: The Missing Links to Information Retrieval[J]. *Computer Journal*, 2017(7):7.
- [7] Losada D E, Parapar J, Barreiro A. A Rank Fusion Approach based on Score Distributions for Prioritizing Relevance Assessments in Information Retrieval Evaluation[J]. *Information Fusion*, 2017, 39:56-71.
- [8] Djenouri Y, Belhadi A, Belkebir R. Bees swarm optimization guided by data mining techniques for document information retrieval[J]. *Expert Systems with Applications*, 2018, 94(MAR.):126-136.
- [9] Losada D E, Parapar J, Barreiro A. Multi-armed bandits for adjudicating documents in pooling-based evaluation of information retrieval systems[J]. *Information Processing & Management*, 2017, 53(5):1005-1025.
- [10] Singh J, Sharan A. A new fuzzy logic-based query expansion model for efficient information retrieval using relevance feedback approach[J]. *Neural Computing & Applications*, 2017,

28(9):2557-2580.

[11] Bo L, Gaussier E, Dan Y. The Dilution/Concentration conditions for cross-language information retrieval models[J]. *Information Processing & Management*, 2018, 54(2):291-302.

[12] Voorhees E M, D Samarov, Soboroff I. Using Replicates in Information Retrieval Evaluation[J]. *ACM Transactions on Information Systems (TOIS)*, 2017, 36(2):1-21.

[13] Chen A R, Chen T, Wang S. Pathidea: Improving Information Retrieval-Based Bug Localization by Re-Constructing Execution Paths Using Logs[J]. *IEEE Transactions on Software Engineering*, 2021, PP(99):1-1.