The Research on personalized Access Control Scheme Based on User Portrait

Jiahui Li
School of Information Management, Beijing Information Science & Technology Beijing, 100192, China; National Engineering Laboratory,Beijing, 100093, China, 17600949929@163.com

Gang Zhao
School of Information Management, Beijing Information Science & Technology Beijing, 100192, China, zhaogang@bistu.edu.cn

Shuai Wang
School of Information Management, Beijing Information Science & Technology Beijing, 100192, China, 782100292@qq.com

Yanzhao Liu
China Information Technology Security Evaluation Center,Beijing, 100085, China; National Engineering Laboratory,Beijing, 100093, China, liuyanzhao@secemail.cn

Jin Chen
China Information Technology Security Evaluation Center,Beijing, 100085, China, chenj@itsec.gov.cn

Follow this and additional works at: https://aisel.aisnet.org/whiceb2019

The Research on personalized Access Control Scheme Based

onUser Portrait

Jiahui Li[1,3] ,Gang Zhao[1] ,Shuai Wang[1] ,Yanzhao Liu[2,3] ,Jin Chen[2]*

[1]School of Information Management, Beijing Information Science & Technology
Beijing, 100192, China
[2]China Information Technology Security Evaluation Center,Beijing, 100085, China
[3]National Engineering Laboratory,Beijing, 100093, China

Abstract: The arrival of the era of big data has made data become an important economic asset, and its security issues have gradually been valued by people. The big data platform usually face many users, traditional access control techniques cannot effectively evaluate whether users are in a trusted state, thus, this paper introduced the concept and method of user portrait and proposed a personalized access control scheme by analyzing the log of hadoop system, tracking user access behavior trajectory, which can protect data security by reasonable authorization.

Keywords: big data, user portrait, access control

1. RESEARCH QUESTION

After an in-depth study of Hadoop ecological principles, this paper considers the following three scenarios are the three most common and urgent security threats in the field of access control.

1.1 Overstepping behavior

The overstepping behavior can be divided into horizontal ultra vires and vertical ultra vires. Horizontal ultra refers to some users try to access other users' resources when they have the same level. Vertical ultra vires means that a user has access to resources beyond his or her own authority.

1.2 System usage exception

When users use Hadoop platform excessively in their normal identity, it indicates that they have more business needs in the near future. If the user overuses Hadoop platform with an abnormal identity, this situation needs to be more vigilant

1.3 The inappropriate authority scope of users

Improper authorization is very common in the actual authorization of big data platform. Some users have a wide range of authority, but the set range of frequently accessed resources is very small.

2. MAJOR RESEARCH FINDINGS

2.1 Overall scheme of the personalized authorization

Users After kerberos certification will leave large number of access records in hadoop platform logs and audit logs after accessing the resources of Hadoop components. This scheme collects users' authority information through the plug-ins of HDFS, HBase and HIVE components in Hadoop platform, and collects log information in Hadoop nodes by using ELK log collection tool. The above information will be stored in the database of the personalized authority control system as the data support for the subsequent establishment of user portraits. Through the above collected data, user portraits are modeled, and the user portrait is obtained as an important basis for judging users' suspiciousness, and the users' authority adjustment proposal will be generated according to the result of user portrait.

[*] Gang Zhao. Email: zhaogang@bistu.edu.cn.17600949929@163.com (Jiahui Li),782100292@qq.com(Shuai Wang) , liuyanzhao@secemail.cn (Yanzhao Liu), chenj@itsec.gov.cn (Jin Chen)

2.2   The construction of the suspicious portrait

User portrait is the core of personalized authorization scheme. The construction process of user portrait is introduced as follows. Firstly, user authorities information and behavior log data can be collected through Hadoop components and ELK tools. In order to make the collected data easy to model construction, it is necessary to do some data preprocessing work such as denoising, integration and data standardization. In order to facilitate business analysis, it is necessary to divide the data into dimensions and design a structured label system based on the specific business status.

The preprocessed data need to be applied to machine learning algorithms or other statistical methods to obtain user tags for Hadoop security scenarios. In order to flexibly expand the acquired tags, these tags need to be stored in the database in a vertical table first, then, in order to facilitate the tag management such as subsequent tag addition, deletion, the tags need to be stored in the database in a horizontal table, finally forming a tag library for the Hadoop security scenario. Eventually, by labeling different tags to different users to build user portraits. According to the portrait, the threat of different users to Hadoop big data platform is distinguished, and the personalized authorization of big data platform is realized.

2.3   The application of user suspicious Portrait

If the status of a user is normal, the existing authority of the user will not be changed, and allowing the user to access resources of Hadoop platform normally. If the user status is abnormal, it is necessary to continue to combine relevant user tags from different dimensions to determine the association rules between the user and resource, so as to obtain user authority adjustment suggestions.

3.   CONCLUSIONS

In view of the deficiency of trusted access control in big data platform, this paper starts from the main security threats in the field of access control, fully considers the subject and object factors in the big data environment, introducing the concept and method of user portrait, has proposed a personalized access control scheme for hadoop system. The administrator can determine whether users have the tendency of unauthorized access, inappropriate scope of authority and excessive usage of hadoop system by observing the change of portrait tags. We hope that this scheme can give full play to its role in protecting data security and promoting the healthy development of business at the level of access control.

ACKNOWLEDGEMENT

REFERENCES

[1] Xinhui Cui, Gang Chen, Zhiqiang He. (2016). Research on Access Control Technology of Cloud Data in Big Data Environment. Modern Electronic Technology, 39(15):67-69, in Chinese.

[2] Sharma DH, CA Dhote, Dr, Potey M M. (2016). Identity and Access Management as Security-as-a-Service from Clouds. Procedia Computer Science, 79:170-174.

[3] PENG J, CHOO K K R, ASHMAN H. (2016). User profiling in intrusion detection:a review. Journal of network&computer applications, 72(1): 14-27.

[4] Juan Tian, Dingwei Zhu, Wenhan Yang. (2018). A review of corporate portraits based on big data platform. Computer Science, 45(S2): 58-62, in Chinese.