

Research on Data Mining Service Mode under the Environment of Cloud Computing

Wei Huang

School of Software Outsourcing, Jishou University, Zhangjiajie, Hunan, China

25752810@qq.com

Keywords: Cloud computing; Data mining; Service; Mode; Exploration

Abstract. Along with the arrival of information era, the network is inundated with all kinds of information, and people are facing various kinds of confusion among the mass information, data analysis becomes a relatively hot potato. The proposal of cloud computing is able to well solve this problem and improve the efficiency of data mining greatly. In order to make people have more profound understanding about the data mining service mode under the environment of cloud computing, and also with the aim of providing certain reference significance for relevant research in future, this paper will combine with the practice of current data mining service, and carry out exploration towards the data mining mode under the environment of cloud computing as well.

Introduction

Nowadays, cloud computing has been mentioned increasingly by people and become the social focus gradually. Particularly for every information technology company, cloud computing has become the core problem in research. By virtue of high-speed broadband and virtualization technology, and meanwhile combining with all kinds of operations, the environment of cloud computing has been formed. Although cloud computing has been mentioned increasingly by people, through making a general survey of the development condition of cloud computing field in our country, the cloud computing power is still at the relatively low stage, and it has not been endowed with definite meaning by its field yet. However, by means of grasping basic knowledge in cloud computing, it is easy to find that it possesses many characteristics, such as large-scale computing power, capacity of resource storage, dynamic shrinkage of data and so on, which also have tremendously improved the possibility that data mining turns to be commercialized under the environment of cloud computing.

Analysis of Data Mining Service Establishment under the Environment of Cloud Computing

What is data mining service? Generally, it refers to the function and behavior gatherings that may involve with data mining. Its fundamental operational procedure is selecting data-pre-processing data-data integration-mining data normally, then reaching the results and conducting evaluation, these functions could be mixed and matched, and thus could be formed new composite application[1]. The computer equipment, server cluster, storage device, development environment and so forth have a shared network environment and rely on it through cloud computing, under such network environment, by virtue of virtualization technology, interface technology, integration technology and others, the software and hardware are packaged and formed corresponding service modules to respond the service requests from three different levels, namely the infrastructure, platform development and application, after the whole process, the complete service mode is taken shape gradually[2]. Under this background, data mining could offer customers integrated capacities required by data development and data application, and thus provide the optimized solution for data mining service. According to the demand and behavior of data mining, and also combine the architecture of cloud computing with SOA framework, the service framework of data mining under the environment of cloud computing has been designed, see Fig. 1.

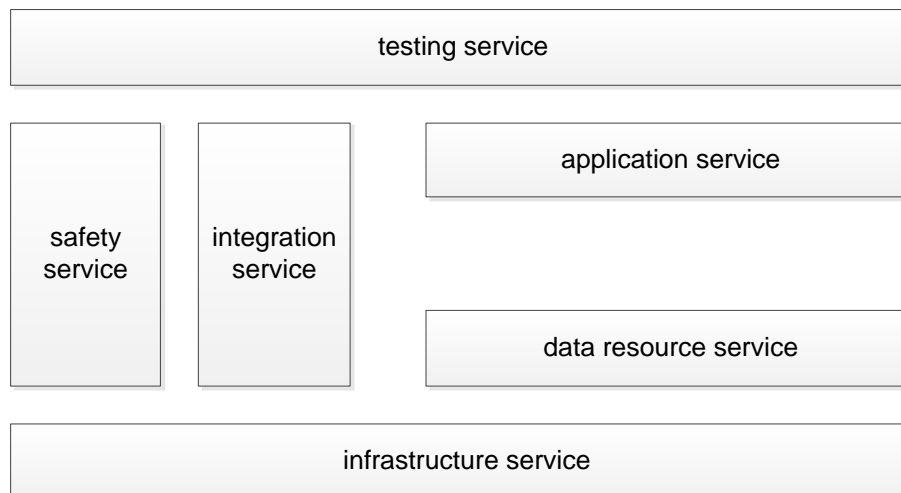


Figure 1. Data Mining Service Architecture Based on Cloud Computing

Seeing from the figure above, we could know that the center of data service is infrastructure service. It provides the computing resource that may be needed for data mining through service mode, and meanwhile provides the ability while accessing these resources remotely. The data resource service provides remote database hosting service, in this way users could use remote data resource more directly and vividly, just like using local database, at the same time, it could provide demand-driven database and data warehouse during accessing. The process service could be offered for data mining business, multi-system spanning could be realized, the data mining module and data information could be bound together, and thus the remote resource for the process could be established[3]. The application service is often called software-as-a-service, it regards the data mining application program as an integral whole, conveys to the end-user via network platform for satisfying the end-user. Under the premises of drawing support from remote hosting testing tools, the testing service could test the mining system in local data[4]. The platform service could offer remote development service for data mining application, to be specific, its service covers the development of application program, data development, interface development and other functions, through which, users could establish enterprise-level application for data mining. The integration service needs to apply abstract interface technology, flow control technology, integrated design and other technologies, so as to offer integrated functions for heterogeneous systems in the data mining application, and then present in front of the users in the form of service[5]. Finally, compared to other services, the safety service is relatively weak, by virtue of the encryption service and security service in data mining, it makes data mining safer under the environment of cloud computing.

Process Analysis of the Establishment of Data Mining Service Mode under the Environment of Cloud Computing

To realize data mining service, the crucial point is to analyze the above-mentioned service framework, and thus to establish the model of service component for every data mining service[6]. Through identifying and describing service, associating data, a complete service component could be constructed, see details in Fig. 2.

Firstly, define service. Through comprehending the information in the problem domain and collecting the relevant information in data mining service, the core function in the corresponding service will be defined, and meanwhile, the required data support, service output and so on will be defined[7]. After defining the service, the description related to data mining service could be obtained, and the candidate service in the service catalogue of data mining could be formed.

Secondly, bind data and service. The composition of data mining service involves with data and function, it refers to establish corresponding relation between candidate service and its associated

data in accordance with the service definition mentioned above. After the entire process, the required information for the service could be found, at the same time, the information call by the service could be stipulated and the service of associated service could be obtained.

Thirdly, establish service component. It refers to integrate the data bound by service and the methods described in the service, and thus to realize the definition function, package the service into independent component, define the uniform interface for service call, and form the independent, complete service component.

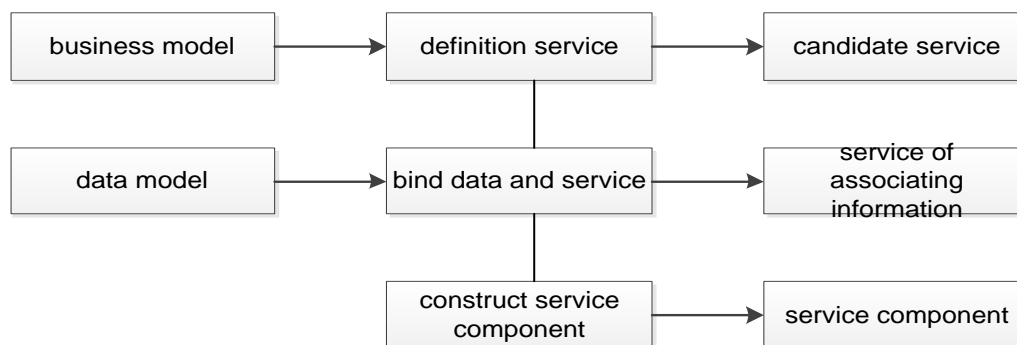


Figure 2. Establishment Process of Data Mining Service

Analysis of the Architecture of Data Mining Service Mode under the Environment of Cloud Computing

The realization of data mining service framework under the background of cloud computing requires 4 levels of structures, see details as below:

Firstly, level of infrastructure. Just as its name implies, it offers the essential storage resource and computing for data mining. In this level, all kinds of resources and networks could be combined via the terminal interface, thus the interconnection and share of physical resource could be realized, and meanwhile the interface could be provided for virtualization process[8].

Secondly, level of virtualization. By virtue of virtualized technologies and tools, it gathers the scattered information resources together under the environment of cloud computing, and packages into unified, transparent service resources, it manages the resources effectively, allocates and dispatches the resources reasonably, and provides for the application and development of platform level[9].

Thirdly, level of platform. The core of the data mining service is the data mining service in level of platform. Such service level is able to provide all kinds of services and functions for the implementation and management of data mining service, such as establishing service catalogue, service combination and so on, it could also be applied to task description, computer resource all during implementing service[10].

Fourthly, level of application. The levels of application interface and application terminal are the constituents in it. The most highlighted feature in the level of application interface is that it could provide management service for authentication of user, user management, demand processing, demand expressing, etc. The level of terminal could offer different access interfaces for different access media. If users want to access and use cloud data, they could realize through portal page, PC or mobile terminal and others[11].

Process Analysis of Data Mining Service under the Environment of Cloud Computing

Under the background of cloud computing, the process of data mining could be divided into five stages, namely analyzing stage, designing stage, developing stage, maintaining stage and declining stage. The analyzing stage requires discovering and recognizing data mining service. On the basis of defining and describing data service, the designing stage designs service interface, terms and

others. The maintaining stage is built up based on interrupting service, it perfects the problems existed during development process. Aiming at the functionality in service, the developing stage forms independent service team. The declining stage refers to when the data service could not satisfy users, the service will be cancelled.

Conclusion

To sum up, the data mining service under the environment of cloud computing is not full-fledged yet, many inadequacies still need to be improved and perfected. Through continuous improvement and perfection, the data mining service could provide higher-quality service for users.

Acknowledgements

A Project Supported by Scientific Research Fund of Hunan Provincial Education Department
Project number: 14C0921

References

- [1] Lanxiang Chen. Using algebraic signatures to check data possession in cloud storage[J]. Future Generation Computer Systems. 2013(7):20-24.
- [2] Yan Zhu, Hongxin Hu, Gail-Joon Ahn, Stephen S. Yau. Efficient audit service outsourcing for data integrity in clouds[J]. The Journal of Systems & Software . 2011 (5):34-38.
- [3] Sangjae Lee, Sungil Jeon, Doyoung Kim. The impact of tour quality and tourist satisfaction on tourist loyalty: The case of Chinese tourists in Korea[J]. Tourism Management. 2010 (5):112-118.
- [4] K Keahey,R Figueiredo,J Fortes. Science Clouds: Early Experiences in Cloud Computing for Scientific Applications[D].Cloud Computing & Its Applications, 2014,17(11):745-748.
- [5] RL Grossman, Y Gu, M Sabala. Compute and Storage Clouds Using Wide Area High Performance Networks[J].Future Generation Computer Systems, 2014, 25(02):179-183.
- [6] R Grossman, Y Gu. Data mining using high performance data clouds: experimental studies using sector and sphere[D]. Acm Sigkdd International Conference on Knowledg,2013,98(07):920-927.
- [7] Jinlin Chen, Keli Xiao. BISC[J]. ACM Transactions on Knowledge Discovery from Data (TKDD). 2010 (3):334-339.
- [8] Chung-Hong Lee, Tzan-Feng Chien. Leveraging microblogging big data with a modified density-based clustering approach for event awareness and topic ranking[J]. Journal of Information Science. 2013 (4):320-328.
- [9] J Wang, J Wan, Z Liu, P Wang.Data Mining of Mass Storage Based on Cloud Computing[D].Ninth International Conference on Grid & Cloud,2013,257(12):426-431.
- [10]G Piatetsky-Shapiro. Knowledge discovery in databases: 10 years after[J].Acm Sigkdd Explorations Newsletter[J].2012,1(02):59-61.