



Efficient Cloud Mining Using RBAC (Role Based Access Control) Concept

Ms. Sunita*, Prachi

CSE, SES, BPSMV

India

Abstract— Cloud computing is an emerging computing paradigm in which resources of the computing infrastructure are provided as services of the internet. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. In this dissertation our main objective is to provide an algorithm with the help of which we can make data secure from unauthorized users. As uses of data come on front we have to face concept of data mining. Data Mining is a field where accuracy matters a lot. The concept behind this dissertation is to provide access control on data by using RBAC (Role Based Access Control). RBAC is called role based access control which would restrict the wrong accessibility of the data from user. No fake submissions would be there and hence the integrity of the data sets which is there in the database will become stronger and hence at the time of the application run, it would give you a better accuracy program. For this purpose we would be using a SVM algorithm and to implement SVM we need an architecture called role based security. Previously in cloud mining no such RBAC concept has been implemented to make the data integrity stronger.

Keywords— cloud computing, data mining, cloud mining, RBAC, SVM (support vector machine).

1. Introduction

The Cloud, as it is often referred to, involves using computing resources – hardware and software – that are delivered as a service over the Internet. Many companies are choosing as an alternative to building their own IT infrastructure to host databases or software, having a third party to host them on its large servers, so the company would have access to its data and software over the Internet. The use of Cloud Computing is gaining popularity due to its mobility, huge availability and low cost. Cloud computing represents both the software and the hardware delivered as services over the Internet. Cloud Computing is a new concept that defines the use of computing as a utility, that has recently attracted significant attention. The deployment models of cloud computing are private Cloud, community cloud, public cloud and hybrid cloud.

Some aspects regarding cloud mining

Cloud mining represents finding useful patterns or trends through large amounts of data. Data mining is defined as a “type of database analysis that attempts to discover useful patterns or relationships in a group of data. The analysis uses advanced statistical methods, such as cluster analysis, and sometimes employs artificial intelligence or neural network techniques. A major goal of cloud mining is to discover previously unknown relationships among the data, especially when the data come from different databases.”

2. Cloud mining techniques

Clustering: Useful for exploring data and finding natural groupings. Members of a cluster are more like each other than they are like members of a different cluster. Common examples include finding new customer segments and life sciences discovery. Classification Most commonly used technique for predicting a specific outcome such as response / noresponse, high / medium / low value customer, likely to buy / not buy.

Association: Find rules associated with frequently cooccurring items, used for market basket analysis, cross-sell, root cause analysis. Useful for product bundling, instore placement, and defect analysis.

Regression: Technique for predicting a continuous numerical outcome such a customer lifetime value, house value, process yield rates.

Attribute Importance: Ranks attributes according to strength of relationship with target attribute. Use cases include finding factors most associated with customers who respond to an offer, factors most associated with healthy patients.

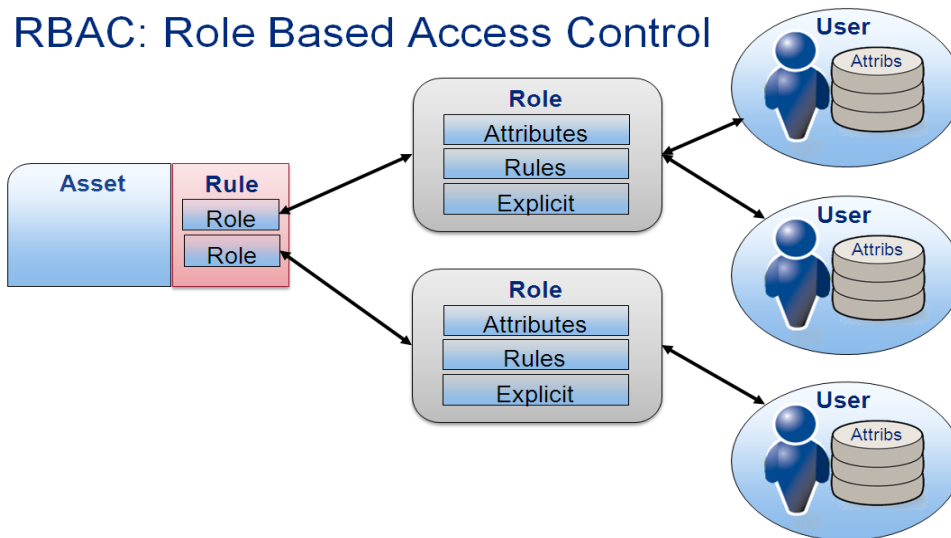
Feature Extraction : Produces new attributes as linear combination of existing attributes. Applicable for text data, latent semantic analysis, data compression, data decomposition and projection, and pattern recognition.

Data mining in Cloud Computing: Data mining techniques and applications are very much needed in the cloud computing paradigm. As cloud computing is penetrating more and more in all ranges of business and scientific computing, it becomes a

great area to be focused by data mining. "Cloud computing denotes the new trend in Internet services that rely on clouds of servers to handle tasks. Data mining in cloud computing is the process of extracting structured information from unstructured or semi-structured web data sources. The data mining in Cloud Computing allows organizations to centralize the management of software and data storage, with assurance of efficient, reliable and secure services for their users." As Cloud computing refers to software and hardware delivered as services over the Internet, in Cloud computing data mining software is also provided in this way.

3. Role-based access control (RBAC)

It is a method of regulating access to computer or network resources based on the roles of individual users within an enterprise. In this context, access is the ability of an individual user to perform a specific task, such as view, create, or modify a file. Roles are defined according to job competency, authority, and responsibility within the enterprise. When properly implemented, RBAC enables users to carry out a wide range of authorized tasks by dynamically regulating their actions according to flexible functions, relationships, and constraints. This is in contrast to conventional methods of access control, which grant or revoke user access on a rigid, object-by-object basis. In RBAC, roles can be easily created, changed, or discontinued as the needs of the enterprise evolve, without having to individually update the privileges for every user



USING ONTOLOGY FOR RBAC Define Roles with Semantic Information. The first question is how to define roles given a specific domain. Ontology, a conceptual structure which contains knowledge in a domain and their relationships, provides useful and valuable information for cloud computing. It specifies a conceptualization of a domain in terms of concepts and their relationships, which is used to generate a commonly agreed vocabulary for information exchange without ambiguity. In general, according to the semantic in a specific domain, roles are defined as a combination of the official positions, job functions, and etc. For example, in an IT company, typical official positions could be that of the ordinary member, group manager, regional manager and etc. Functions represent the user's daily duties such as being a developer, testing engineer and etc. Additionally the organizational unit to which a user belongs is used as an access control criterion for certain applications. All these data is defined and maintained in the human resources (HR) database. Thus, a RBAC system has an accurate image of the current organizational status and existing roles. A sample role definition is shown in Table I, e.g. CEO has a unique role ID = 3. Note that each employee can be assigned to one or more roles. Manage Roles Hierarchy with Ontology In a domain, multiple possible role hierarchies are defined by ontology systems from different communities. Examples of possible role hierarchies in an IT company are shown in Existing research have discussed how to define an ontology in a specific domain, compare and integrate ontology systems between different communities. In practices, for a given domain, multiple reference ontology systems from various communities may in that domain. Solve Role Hierarchy using Ontology Trees Role hierarchies impose restrictions which can generate a simpler tree structure (i.e., a role may have one or more immediate ascendants, but is restricted to a single immediate descendant). To extend the traditional definition of trees for an ontology in a specific domain. For doing cloud mining using RBAC we have used SVM algo. By using RBAC we are able to control access of data by the users.

4. Conclusions

In cloud computing, security is an important issue due to the increasing scale of users. Current approaches to access control on clouds do not scale well to multi-tenancy requirements because they are mostly based on individual user IDs at different granularity levels. However, the number of users can be enormous and causes significant overhead in managing security. RBAC (Role-Based Access Control) is attractive because the number of roles is significantly less, and users can be

classified according to their roles. This paper proposes a RBAC model using a role ontology for Multi-Tenancy Architecture (MTA) in clouds. The ontology is used to build up the role hierarchy for a specific domain. Ontology transformation operations algorithms are provided to compare the similarity of different ontology. The proposed framework can ease the design of security system in cloud and reduce the complexity of system design and implementation.

Acknowledgment

I am thankful to my guide Ms. Sunita Assistant. Professor in Department of Computer Engineering in B.P.S.M.V Khanpur, Sonepat for their valuable support.

References

- [1] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, *Above the Clouds : A Berkeley View of Cloud Computing*, 2009.
- [2] I. Foster and C. Kesselman, "The grid: blueprint for a new computing infrastructure," Oct. 1998.
- [3] S. Clearwater, *Market-Based Control: A Paradigm for Distributed Resource Allocation*, World Scientific, 1996.
- [4] J. Shneidman, C. Ng, D.C. Parkes, A. AuYoung, A.C. Snoeren, A. Vahdat, and B. Chun, "Why Markets Could (But Don't Currently) Solve Resource Allocation Problems in Systems," *Challenges*, 2005, p. 7.
- [5] P. Cramton, Y. Shoham, and R. Steinberg, *Combinatorial Auctions*, The MIT Press, 2005.
- [6] A. Das and D. Grosu, "Combinatorial auction-based protocols for resource allocation in grids," *Parallel and Distributed Processing Symposium, 2005. Proceedings. 19th IEEE International*, 2005.
- [7] A. Marshall, *Principles of Economics (Great Minds Series)*, Prometheus Books, 1997.
- [8] "Market," *From Wikipedia, the free encyclopedia* <http://en.wikipedia.org/wiki/Market>. 113 |
- [9] D. Friedman, "The double auction market institution: A survey," *The Double Auction Market: Institutions, Theories, And Evidence*, J. Rust, ed., Westview Press, 1993, pp. 3-26.
- [10] R. Buyya, D. Abramson, J. Giddy, and H. Stockinger, "Economic models for resource management and scheduling in Grid computing," *Concurrency and Computation: Practice and Experience*, vol. 14, 2002, pp. 1507-1542.
- [11] Y. Shoham and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*, Cambridge Univ Pr, 2009.
- [12] D.S. Diamond and L.L. Selwyn, "Considerations for computer utility pricing policies," *Proceedings of the 1968 23rd ACM national conference*, New York, New York, USA: ACM Press, 1968, pp. 189-200.
- [13] I.E. Sutherland, "A futures market in computer time," *Communications of the ACM*, vol. 11, Jun. 1968, pp. 449-451.
- [14] R. Buyya, D. Abramson, and S. Venugopal, "The Grid Economy," *Proceeding of the IEEE*, vol. 93, Mar. 2005, pp. 698-714.
- [15] C.S. Yeo and R. Buyya, "A taxonomy of market-based resource management systems for utility-driven cluster computing," *Software: Practice and Experience*, vol. 36, Nov. 2006, pp. 1381-1419.
- [16] R. Buyya and M. Murshed, "GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing," *Concurrency and Computation: Practice and Experience*, vol. 14, Nov. 2002, pp. 1175-1220.