Rough Set Theory Based Technique For Detection Of Anonymous Cloud Users Without Trusted Third Party

Smt. K. Venkata Ramana,
Associate Professor,
Department of CSE,
RVR&JC College of Engineering,Guntur, India.

Dr. B. Raveendra Babu,
Dean, Administration,
VNR Vignana Jyothi Institute of Engineering &Technology, Hyderabad.

Abstract

The cloud service authentication is significantly important to obtain a secure access from the service provider or doing the operations in cloud data. This access can cause security problems since there may be possibility that the anonymous users can get access to do the operation over the cloud data like credential users. So, there is a need of anonymous credential systems without relying on trusted third party. These misbehaving activities should be avoided with the help of effective blacklist able mechanism to block the entire anonymous cloud users. With the above objectives, the anonymous credential system will be developed for cloud users to provide the privacy for users. The main advantages of the proposed system are that it can preserve their anonymity, block the intruders and that to without using Trusted Third Party (TTP). By taking all the advantages, the cloud computing-dependent technique will be developed based on Blacklist able Anonymous Credentials (BLAC) technique without TTP. In the proposed technique, rough set theory-based rules are utilized to block the misbehaving users. Rough set theory can be regarded as a mathematical tool for the analysis of imperfect data. On considering the data, which is given as input, it can be seen that the proportions of blacklist in different set of users are also equitant. Proposed approach is found to be efficient in identifying the blacklists.

Keywords: *Cloud Security, BLAC, Rough Set theory, Trusted Third Party*

1. Introduction

The recent cloud technologies depend on value of the world’s norms, the ease of use by end users and most significantly the grade of information security and control. Cloud computing becomes a new and evolving technology, which alters the way IT architectural solutions are given forward by means of touching the theme of virtualization: of data storage, of local networks (infrastructure) as well as software [8,7]. Cloud computing is evolving as the advanced distributed computing standard and is increasing the interests of researchers in the field of Distributed and Parallel Computing [2], Service Oriented Computing [3] and Software Engineering [4]. In Cloud computing which uses virtual machines, physical change or migration may not alter the services provided by the service provider. More services are provided to the user as per the user demand without having any concern with the physical hardware [5].

The virtual server from the logical server set comes up with a number of security issues [9]. Cloud computing poses various challenges in security threats. Under cloud computing, conventional cryptographic primitives will not be selected directly for security because user’s lack control of data. The proof of corrected data storage in the cloud should be executed without clear information of the complete data. The long term unceasing assurance of the data safety for different users is even more challenging. Cloud Computing is not just a third party data warehouse. The data in the cloud can be restructured by users by making insertion, deletion, modification, appending, reordering, etc. [6].

In cloud computing, when anonymous access to service providers (SPs) is considered, it provides users a major level of privacy and it could assure users the license to create troubles without bothering about punishment. For example, Wikipedia provides editors to update content anonymously and that results in several users to misbehave by uploading irrelevant content. The cloud services eventually expect a level of responsibility over misbehaving users. Numerous anonymous credential systems have been proposed in which users can be efficiently deanonymized [10]. Usually, an anonymous credential system associates two needful components –protocol to get signature on user's personal key and monogram on user's personal key. Signature is obtained from the authority without losing any useful information [12]. In cloud computing environment, most of the anonymous credential systems recently proposed is web-based and there are options of updating the database by different users.

.

In cloud computing, there is no information available about where the data is stored, who has access to them or any data transferred [10]. Methods that use a TTP increase the risk of attacks like correlation attacks which happen when an entity acquires a set of multiple data and is able to correlate it to the physical identity of an entity such as a person. Methods that do not use TTP reduce such attacks. It is also risky in assuming that the TTP will act as it is expected, which may not be true all the time. The proposed approach is independent of the usage of TTP.

In this work, a method based on the rough set theory without TTP is proposed to handle the activities of anonymous users in the cloud network. Rough set theory is used to analyze imperfect data. Any set of all indiscernible (similar) objects is called an elementary set, and forms a basic granule (atom) of knowledge about the universe. Any union of some elementary sets is referred to as a crisp (precise) set – otherwise the set is rough (imprecise, vague). The users usually make use of different cloud resources and that may result in unexpected behaviour in the cloud resources. Sometimes users may update the resource with irrelevant contents that may shoot as a problem. In order to tackle these problems, an efficient user activity mapping method is needed. Thus a rough set theory based method is proposed to evaluate user in a cloud network without any trusted third party. The proposed approach mainly includes three phases, keeping the user anonymous, cloud resource authentication and user activity validation. Experimentations are done to evaluate the proposed method.

The main contributions of the approach are,

- The user validation is done without using a TTP
- A rough set theory is used to validate user activities

The rest of the paper is organised with different sections like literature survey, proposed methodology, experimental analysis and conclusion.

2. Review of related works

Jan Camenisch and Anna Lysyanskaya [1] have proposed a method for non-transferable anonymous credentials with optional anonymity revocation. It was implemented by a primitive, known as circular encryption. It was a practical anonymous credential system that is based on the strong RSA assumption and the decisional Diffie-Hellman assumption modulo a safe prime product. It is a useful solution that permits a user to anonymously exhibit possession of a credential, number of times necessary without being caught by the issuing organization. . In order to avoid the misuse of anonymity, their arrangement was to first offer possible anonymity revocation for selected transactions. This method prevents users from sharing their credentials by using all-or-nothing sharing, which is a less efficient mechanism because the organizations have to store 25K bits per user when using RSA moduli of length 1024 bits.

MalikaIzabachene et al [12] has presented an anonymous credential scheme with non-interactive proofs of credential possession where credentials were associated with a number of attributes. The recent outcomes of Camelish and Gro(CCS 2008), proved and assures the verifier that qualified attributes can satisfy a certain predicate. The creation of them relies on a particular type of P-signature, termed block-wise P-signature that permits a user to be gained a signature on a committed vector of messages and allows it possible to be generated a short witness that serves as a proof that the signed vector satisfied the predicate. This method cannot prevent the public key size from depending on n number of attributes.

Rohit Ranchal et al [10] has proposed a method which was independent of TTP and used to identity data on untreated hosts. Their method was formulated on the predicates over encrypted data and multi-party computing for negotiating an activity over a cloud service. The method possess an active bundle, which has a middleware agent that includes PII data, privacy policies, a virtual machine that enforces the policies, and has a set of protection mechanisms to defend itself. An active bundle is used on behalf of a user to be authenticated to cloud services for user's privacy policies. This method is prone to attacks like the active bundle may also be not executed at all at the hosts of the requested service. In this case its data is not disclosed but the user is denied access to the service that he requests.

Patrick P et al [11] have proposed a method in which service providers can revoke the credentials of misbehaving users without relying on a TTP. It supports a n-strikes-out revocation policy, whereby users who have been consecutively misbehaved at least n times are revoked from the system. This method lacks efficient computation at the service provider that is linear to the size of the blacklist. This method does not consider multiple combinations of misbehaviour such as "user has defaced a webpage AND user has posted copyrighted material".

3. Rough Set Theory based Technique for Detection of Anonymous Cloud Users without Trusted Third Party

The cloud system is driving an explosion of innovation with tremendous social, scientific, and business benefits in the recent times. As the traffic through cloud network increases, similar to every other network, the cloud networks also are affected by false authentication. The failures in the authentication triggers problems in cloud resources and which may result in the total system failure. The proposed approach produces a system for identifying the anonymous users in the cloud network. A rough set theory based algorithm is proposed here to identify the actions of anonymous users and to identify their validity in the network. The proposed approach is executed as a procedure manner and includes many steps to identify the user's credibility in the cloud network. The users, who make good and valid updates in the cloud resources, are accepted by the network and the users, who alter the cloud resources, are denied and black listed by the cloud system. The major steps of the proposed approach are given below,

a) Preserving the anonymity of users
b) Cloud network authentication
c) User activity validation

The above three processes are considered as the major steps in the detection of anonymous user's activity in the cloud network. One of the major specialities of the proposed approach is, there is no need of any kind of third party resources for validation. The proposed rough set theory based algorithm triggers the validation process, which can be considered as advantage to the proposed approach as minimum load is engaged in the cloud network for authentication purpose.

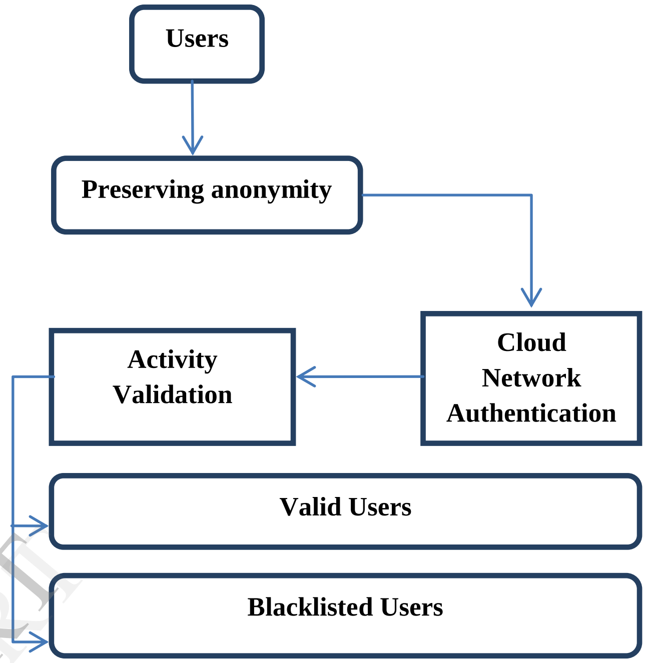


Figure 1. Overview of the proposed approach

The figure 1 presents the overview of the proposed rough set theory based method to identify the anonymous user's activity in the cloud network. Let us consider the proposed approach in detail:

3.1. Preserving the anonymity of users

The cloud system can be accessible to any user, who is having a possible resource to access the cloud network. The cloud network with anonymous user is considered to preserve confidentiality. When a user accesses the cloud resource, they may alter the resources, so the user should be identified in order to record their edits. Thus for this purpose, a certificate authentication system is used, that will provide a unique identity to each user. For each of the user in order to attain the certificate of authorization from the certifying authority, a request is processed and sent to the certifying authority. The system that provides the certificate issues an anonymous certificate to each of the user according to the request provided. The fields that the user has to input consist of name *N*, age *A*, address Add, phone number *Ph*, identification *ID*, and mail *mID*. When we consider the ith user on the cloud network is considered, the fields are represented as N_i, A_i, Ph_i, ID_i, Add_i *and* mID_i. In-order to avoid any Sybil attack, where a user can register even after being black-listed by giving new values in the respective fields, the paper incorporates the passport number field. As any user will have only one unique passport number, the blacklisted user will not be able to log in and will be shown the message that the user is not permitted to log in and that he/she is in the blacklist. When a user input details, the certificate authority will check for the details provided by the user with data in the blacklist database. If a positive result has been obtained, then the user will not be granted a certificate. The fields, ID and Add are selected as key elements for the blacklist check and the credentials stored in both blacklist table and the system table.

$$user\ i = [N_i, A_i, Ph_i, ID_i, Add_i, mID_i]$$
$$select\ blacklist\_DB(bl\_list\_DB)$$
$$Check\,(ID_i, Add_i)\ with\ (bl\_list\_DB)$$
$$if\ true \rightarrow certificat\,e\ denied$$
$$if\ \ false \rightarrow certificat\,e\ granted$$

With the user input details, the user acquires a certificate in the anonymous form, only if the user is not in the blacklist. The anonymous certificate is provided with details like,

Code Name (CN)	Public Key (P_key)	CA Signature (CAS)	Unique ID (UID)

The CN represents the user pseudo name to ensure the anonymity, a public key P_key is issued for the user to access the CA, each user is provided by signature from the certifying authority (CA) and a serial number provided by the CA is considered as the UID, a unique identifier. Here, the public key is generated by the RSA algorithm whereas pseudo name and signature is obtained using SHA-256 algorithm. Normal certificate will contain all the details of the user along with public key, signature and certificate serial number. Once the certificate is obtained, the user will be able to use this certificate while logging in with the cloud service provider. All the details are fed as input by the user and the details about the certificate issued are stored in the database.

RSA algorithm: RSA algorithm is used for user public-key generation in the proposed method. RSA is a public key cryptographic algorithm. The procedure involved in RSA algorithm for generating public key for the respective user is given below:

- Initially, select two prime numbers *b* and *c* randomly preferably having the same length.
- Compute the modulus *z* for public key given by: z = a × b
- Compute Euler's totient function $\delta(z) = (a - 1) \times (b - 1)$
- Choose an integer *t* such that $1 < t < \delta(z)$ and greatest common divisor of t and δ(z) is 1 and δ(z) is co-prime. And *t* is the public key component.
- The public key consists of the modulus *z* and the public exponent *t*.
- The user name *N* is hashed to integer value *h*.
- With the aid of *z, h and t* final public key is generated as $K = h^t$ (mod z).

SHA 256 algorithm: The Secure Hash Algorithm (SHA) is one of cryptographic hash functions and is used to generate the name (pseudo name) and the corresponding signature of the Certifying Authority (CA). The procedure involved is explained below.

- The field is first padded with its length in such a way that the result is a multiple of 512 bits long.
- Subsequently, it is parsed into 512-bit message blocks $F^1, F^2, \ldots, F^n$

The message blocks are then processed one at a time, beginning with initial hash value H^0, sequentially compute: H i = H (i-1) + L (Fi). (H (i-1)), where L is the SHA-256 compression function, + means word-wise mod 232 addition and H^n is the hash of F.

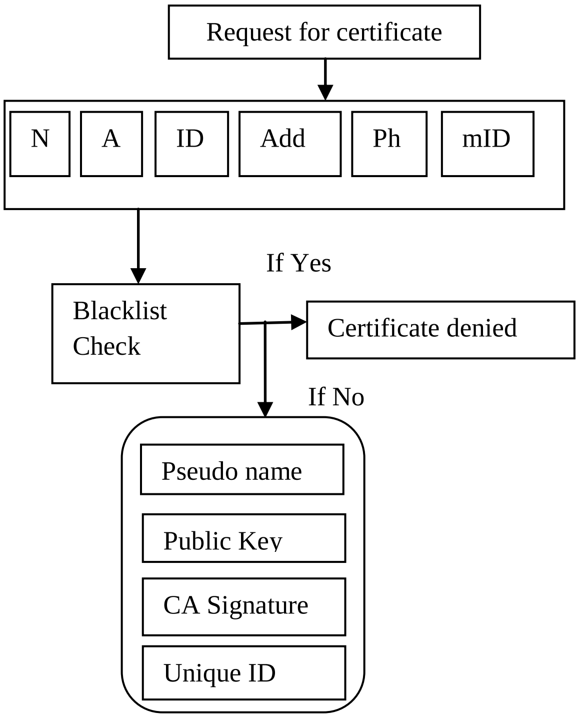


Figure.2. Certificate authentication process

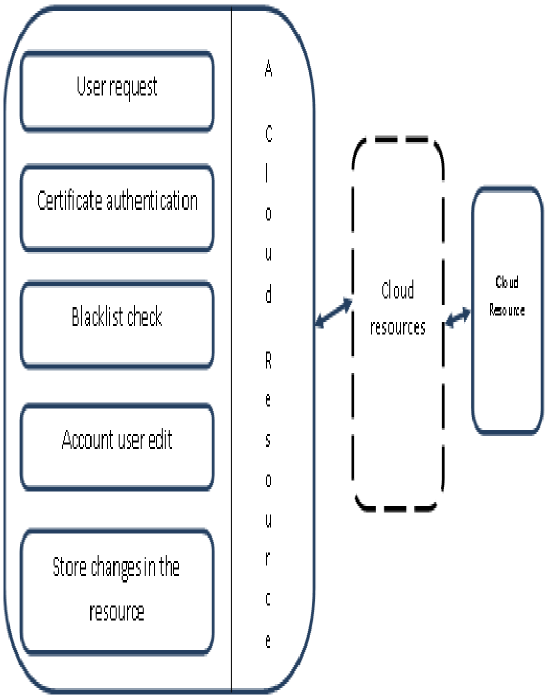


Figure.3. Cloud resources editing

The figure 2 presents the processing of certificate authentication module of the proposed approach. The module is mainly used to create a unique identity to the user, who accesses the cloud network. The module also preserves anonymity of the user at the same time.

3.2. Cloud network authentication

The second major step included in the proposed approach is to avail authentication to the user, who accesses the cloud network. As the authentication is availed, the user activity can be mapped with the help of network traffic. The user initially requests an access to the cloud resources through a cloud resource provider. Upon receiving the request, the user is re-directed to the corresponding cloud resource of the cloud network. Suppose, there are n number of cloud resources and they are represented by R={r1, r2,...,rn}, the user *i* requests any one of the cloud resource $r_{j,}$ where $0 < j \leq n$ and consequently is redirected to the requested resource r_j by the cloud. The user provides the certificate details like username and password and accesses the resource of the cloud resource provider. The following diagram shows the cloud resource access module.

The figure 3 represents the processing that happens in a single cloud resource. Once the cloud resources access is given to the user, the user can perform edits and validation of the contents in the cloud resource. Here, without entering the full personal details, the user can directly access the service provider. In this case, the user makes use of the certificate details for the entry. The user needs to enter the details like,

Code Name (CN)	Public Key (P_key)	Unique ID (UID)	User ID (ID)	Password (W)

When the user provide user ID and password W, the credentials from the certificate are updated on the cloud resource database. Before granting access, the cloud resource provider (CSP) checks if the user is a blacklisted one and if so, the user is denied access. CSP makes use of user ID and unique ID to check if the person is blacklisted. For a user i, it will have the unique ID UID_i and user ID ID_i which will be checked with *UID* and *ID* fields respectively in the blacklist database.

$$\begin{aligned}
&user\ i = [CN_i, P\_key_i, ID_i, UID_i, W_i] \\
&select\ blacklist\_DB(bl\_list\_DB) \\
&Check\,(ID_i, UID_i)\ with\ (bl\_list\_DB) \\
&if\ true \rightarrow User\ Access\ denied \\
&if\ \ false \rightarrow User\ access\ granted
\end{aligned}$$

Thus based on the access level, the user can do the permitted edits and can review the edits of the other user. The modules in the cloud resource map the user’s activity and store it on a table. Once the user session is over, a validation from the system is also provided and stored in the table. The user who got access to the cloud resource can perform edits as well as review the updates by other users. The user review is mapped manually and the user can manually enter the review about the contents to the review table. But in case of system review, the cases are little bit different as a text comparison is engaged to record the user edits. If comparison is positive then updated text is irrelevant and if comparison is negative then updated text is relevant.

3.3. Rough Set theory for activity detection

The most important phase in the proposed approach is to identify the activities of the anonymous users. The updates done by the anonymous users are taken into account and through analysis; the contents are identified as relevant or irrelevant. The level of irrelevance of the content provided by the user is used for updating the black list. The proposed approach uses a rough set theory to identify the level of updates done by the user. The rough set theory is fed with two tables from the proposed approach. The tables are listed as, user review table and system review table. The user review table is generated by giving values to the updates of the user through discretization methods. In discretization, intervals between levels of updates on the contents are identified and a value is given for each level. The same procedure is used for creating the system review table. Consider the following table,

Table.1. Discretization of updated contents

Update by users	User review	System review
Higher relevance	$\mathrm{UR}_{\mathrm{high}}$	$\mathrm{SR}_{\mathrm{high}}$
Lower relevance	$\mathrm{UR}_{\mathrm{med}}$	$\mathrm{SR}_{\mathrm{med}}$
Irrelevance	$\mathrm{UR}_{\mathrm{low}}$	$\mathrm{SR}_{\mathrm{low}}$
Higher irrelevance	$\mathrm{URb}_{\_\mathrm{list}}$	$\mathrm{SRb}_{\_\mathrm{list}}$

The discretization process is identified by calculating deviation between the maximum level of the update to the minimum level and mid-level of the update on the contents. Once the discretization is over, the rough set theory defined methods generate rules based on the discretization values. Thus, when a set of users accessing the cloud network is considered, they are given values according to the above table based on the relevance of their contents they updated. Now, in order to generate the rules based on the rough set theory, the detection method needs to find the similarity between each attribute. In the current context, a user ID is provided with two different attributes, the user review and the system review. The similarity calculation is executed by comparing user review values and system review values of a particular user. The similarity between those values deducts a decision on the user’s credibility. For example, consider user x, if user x possess URhigh and SRhigh then decision on user x is High. Thus, by comparing all the values of every user, a set of decisions based on the similarity of attributes is generated. The fields on the values are defined as, high (H), medium (M), low (L) and very low (VL). Now the similarity values are then taken into account for rule generation. A similarity table is generated for identifying the decision values on each user.

Table.2. Decision values

user	UR	SR	decision
UID1	URb_list	SRb_list	VL
UID2	URlow	SRhigh	M
UID3	URhigh	SRmed	M
UID4	URlow	SRlow	L

Now, based on the decision values, the rough set theory defines rule sets to decide the credibility of the user. The decision values are taken into the account to generate the rules. Each rule is associated with fields from the decision table and each rule is associated with two fields “YES” and “NO”. The value YES triggers the user as an entry to the blacklist and the value NO triggers normal user. The rules can be given as,

$$\begin{aligned}
&Rule1 \Rightarrow if\,(UR \cap SR) = VL, \\
&then\ decision\ is\ "YES" \\
&Rule2 \Rightarrow if\,(UR \cap SR) = L \,||\, (UR \cap SR) = \\
&M \,||\, (UR \cap SR) = H, \\
&then\ decision\ is\ "\,NO"
\end{aligned}$$

According to the decision returned from the rough set theory, the proposed approach put the user into either blacklist database or to the regular database as a valid user in the cloud network.

4. Experimental results and analysis

The rough set theory based activity detection of anonymous users in the cloud network having discussed, the current section includes the experimental analysis of the proposed approach under certain test criteria is considered.

Experimental setup

The proposed technique is implemented in java programming language with JDK 1.7.0 and CloudSim toolkit and the performance of the proposed algorithm is analyzed. For the purpose of evaluation, the details of about 50 users including the user details an the edits done by users are taken. Every time when a user accesses the cloud network by giving either user details, the user edit is made and is reviewed by the certifying authority and other users. Based on above edited data user ids are blacklisted. The results are discussed in the following section.

4.1 Results

The proposed approach is designed with three cloud resources namely, data mining, image processing and networking. The user requests for any one of the cloud resources and once the access is granted by the cloud network, the user can perform edits and reviews. Consider the following output screen, figure 4 represents the cloud resource selection page, from here the user can select their appropriate resource and can request for the access. The page also represents three buttons, in which the user button is used for getting access to edit the contents. The reviewer button is used to review the contents edited by the existing user and the manager button triggers the system review of the contents. The following figure 5, represents the interface of the cloud resource, where user can review and edit the contents.

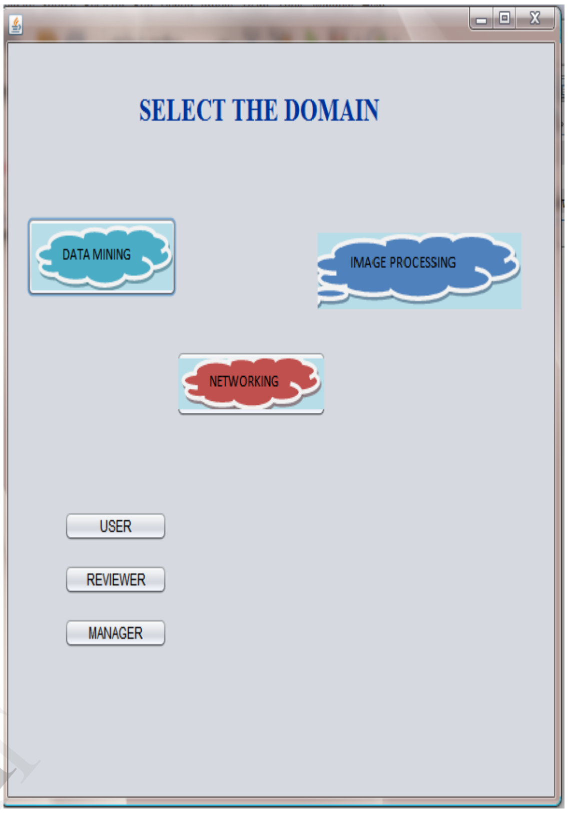
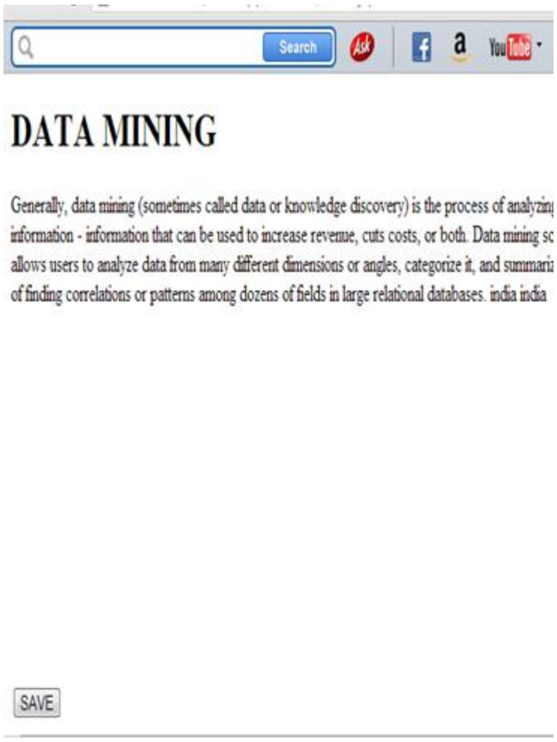


Figure.4. Resource selection

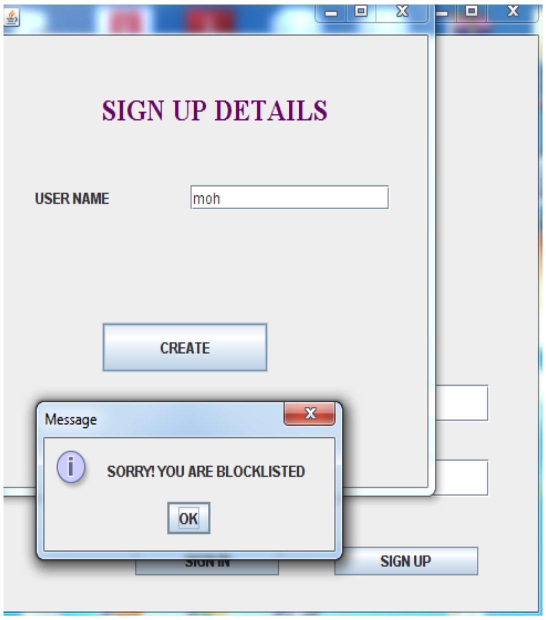


DATA MINING

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzin
information - information that can be used to increase revenue, cuts costs, or both. Data mining sc
allows users to analyze data from many different dimensions or angles, categorize it, and summari
of finding correlations or patterns among dozens of fields in large relational databases. india india

Figure.5. Resource editing section

The figure 6 represents the blacklist check page.



SIGN UP DETAILS

USER NAME: moh

Message

SORRY! YOU ARE BLOCKLISTED

Figure.6. Black list validation

Performance evaluation

Performance analysis of the proposed approach is conducted by selecting 50 different users. The users are selected as different sets of 30, 40 and 50. The responses of different sets are mapped to evaluate the performance of the proposed approach.The performance is evaluated for the parameter, the number of users blacklisted.

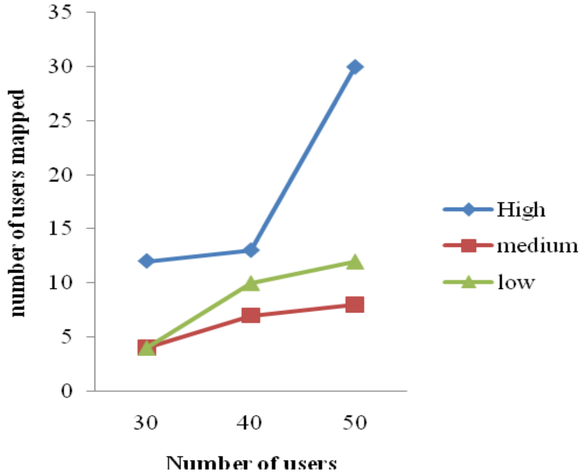


Figure.7. Mapping of decision values based on user review

The figure 7 represents updates of the different set of users mapped based on their level of relevance. The level of relevance can be listed as high, low and medium. The analysis from the figure shows that most of the contents updated by the user are relevant

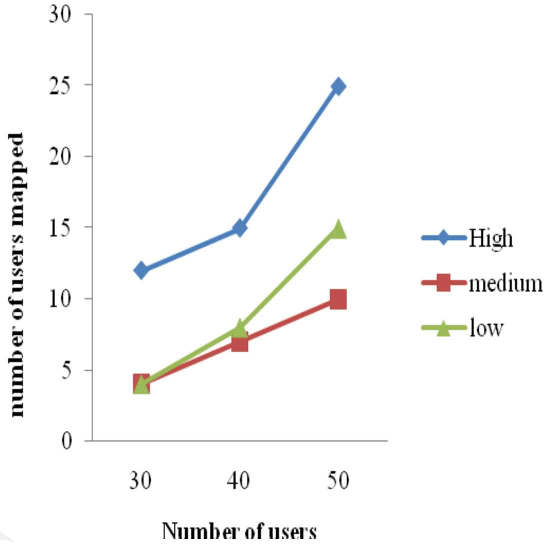


Figure. 8 Mapping of decision values based on system review

The figure 8 represents the mapping of number of users based on the relevance of their edits being validated by the system. The system review analysis also provided the same information that most of the users are updating the cloud resource with relevant information.

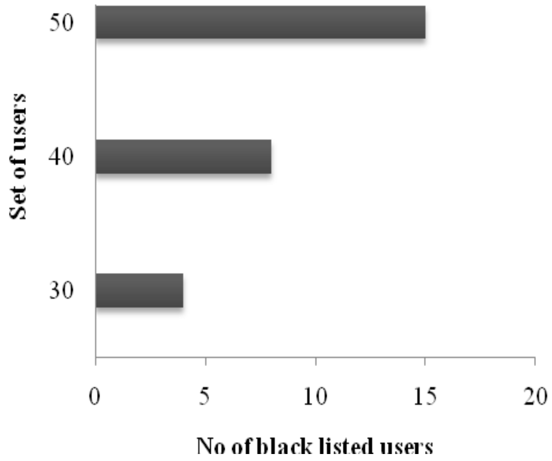


Figure.9. Blacklist mappings

In figure 9, the mappings of the number of blacklisted users in different sets of users, who have accessed the cloud network, are represented. The analysis shows that as the number of users increase, the numbers of blacklist entry also increases. On considering the data, which is given as input, it can be seen that the proportions of blacklist in different set of users are also equitant. Thus, it can be stated that the proposed approach is efficient in identifying the blacklists.

5. Conclusion

The proposed approach provides the anonymous credential system for cloud users to provide the privacy for users. The main advantages of the proposed system is that it can preserve their anonymity, block the misbehaviour users that to without using TTP. By taking all the advantages, the cloud computing-dependent technique is developed based on Blacklist able Anonymous Credentials (BLAC) technique without Trusted Third Party (TTP). In the proposed technique, rough set theory-based rules are utilized to block the misbehaving users.The experimental analyses showed that the proposed approach is good in mapping the activities of the anonymous users. On considering the data, which is given as input, it can be seen that the proportions of blacklist in different set of users are also equitant. Thus, it can be concluded that the proposed approach is efficient in identifying the blacklists.

6. References

[1] Jan Camenisch and Anna Lysyanskaya, “An Efficient System for Non-transferable Anonymous Credentials with Optional Anonymity Revocation", Theory and Application of Cryptographic Techniques: Advances in Cryptology, vol. 3, no.3, p p. 93-118, 2001.

[2] B. Raghavan, S. Ramabhadran, K. Yocum, A. C. Snoeren, "Cloud Control with Distributed Rate Limiting," Proc. 2007 ACM SIGCOMM, pp. 337-348, 2007.

[3] D. Ardagna, B. Pernici, "Adaptive Service Composition in Flexible Processes," IEEE Trans. on Software Engineering, vol. 33, no. 6, pp. 369-384, 2007.

[4] SECES, Proc. First International Workshop on Software Engineering for Computational Science and Engineering, in conjuction with the 30th International Conference on Software Engineering (ICSE2008), Leipzig, Germany, May, 2008.

[5] Dai Yuefa, Wu Bo, GuYaqiang, Zhang Quan, Tang Chaojing, "Data Security Model for Cloud Computing,” Proceedings of the 2009 International Workshop on Information Security and Application, 2009.

[6] Cong Wang, Ian Wang, KuiRen, Wenjing Lou, “Ensuring Data Storage Security in Cloud Computing,"17th International Workshop on Quality of Service, pp.1-9, 2009.

[7] Weinhardt C, Anandasivam A, Blau B, Stosser J, “Business Models in the Service World”, IT Professional, vol. 11, pp. 28-33, 2009.

[8] Leavitt N, “Is Cloud Computing Really Ready for Prime Time?,” Computer, Vol. 42, pp. 15-20, 2009.

[9] Kevin Hamlen, Murat Kantarcioglu, Latifur Khan, BhavaniThuraisingham, "Security Issues for Cloud Computing,” International Journal of Information Security and Privacy, vol.4, no.2, pp. 39-51, 2010.

[10] RohitRanchal, Bharat Bhargava, Lotfi Ben Othmane, LeszekLilien,Anya Kim,Myong Kang and Mark Linderman, “Protection of Identity Information in Cloud Computing without Trusted Third Party", IEEE International Symposium on Reliable Distrubuted Systems, vol. 19, no.5, 2010.

[11] Patrick P. Tsang, Manhoau, Apukapadia and Sean W. Smith, “BLAC: Revoking Repeatedly Misbehaving Anonymous Users Without Relying on TTPs", ACM Transactions on Information and System Security (ACM TISSEC), vol. 13. no. 4,2010.

[12] MalikaIzabachene, BenoitLibert, and Damien Vergnaud, "Block-wise P-Signatures and Non-InteractiveAnonymous Credentials with Efficient Attributes", in Proceedings of the 13th IMA international conference on Cryptography and Coding, pp. 431-45, 2011.