ORPHANAGE HOME MANAGEMENT SYSTEM USING CLOUD WITH DATA ANONYMIZATION

Mr.N.Jayapandian[1], Dr.A.M.J.Md.Zubair Rahman[2], A.Sowntharya[3,] U.Kasthuri[4], M.Sowntharya[5], V.Nivedha[6]

[1]Assistant Professor & Knowledge Institute of Technology, Salem
[2]The Principal & Al-Ameen Engineering College, Erode.
[3,4,5,6]*UG Students & Knowledge Institute of Technology, Salem*

[1]njayapandian@gmail.com
[3]sowntharyaadhi@gmail.com
[4]nivivedhu2@gmail.com

ABSTRACT:

Cloud computing is fast growing a popular option for renting of computing and storage infrastructure services. Cloud computing is a computing resource in which tasks are assigned to a different set of connections, services and software that can be accessed over internet. This paper proposes the usage of cloud computing in Orphanage Home Management System. The main aim of this project is to protect sensitive data of orphans. The Orphan Home Management System is going to be developed to overcome the problems that occur in the center's management. Cloud gives more storage area to the system. The cloud computing relies on the internet. We used data anonymization to enhance security in cloud.

Keywords: Cloud Computing, Orphanage management system, Anonymization.

I. INTRODUCTION

The proposed Orphan Home Management System is expected to overcome the general problems in handling data such as data redundancy, security of data, time consuming and recovery manners.The main aim of our system is to provide high security internet storage. Our system is implemented using cloud computing and Anonymization. Anonymization is a technique that enterprise can use to increase the security of the data in cloud. Data Anonymization is a process of changing the data that will be used or published in a way that prevents the identification of key information. Using data anonymization key pieces of confidential data are obscured in the way that maintains data privacy. For example, If someone sends a file, there may be information on the file that leaves a trail to the sender. The sender's information may be traced from the data logged after the file is sent.

However, once the file is anonymized, data associated with it being sent cannot be traced to the sender, at least in theory-sensitive data sets on cloud probably leads to privacy concerns because of multi-tenancy system. Data encryption and anonymization is two widely-adopted ways to combat privacy breach[2]. The encryption is not suitable for data that are processed and shared frequently and the anonymizing big data and manages anonymized data sets are still challenges for traditional anonymization approaches Existing technical approaches for preserving the privacy of data sets stored in cloud mainly include encryption and anonymization. The cloud computing thus can be used in orphan home management system. Cloud computation provides massive power and storage [1] . Cloud computing refers to the next evolution of the Internet. Instead of buying software, installing it on your computer, upgrading it periodically and storing all your data on your hard drive, with cloud computing you use software applications online, as a service. All you need is your computing device and an Internet connection. The cloud computing thus can be used in orphan home management system.

II.LITERATURE SURVEY

A.Cloud Security

The aspect of security and confidentiality must intervene to protect the data from each of the enterprises. Secure storage and data transmission requires using a modern aspect of encrypting that has the criteria for treatment such as, the necessary time to respond to any request sent from the client and the size of an encrypted data which will be stored and transmitted on the Cloud server. Security has emerged as the most important barrier to faster and widespread adoption of virtualization as well as cloud computing. Security depends from person to person as well as industry to industry how they analyze the concept of security in Cloud Computing [7].

B. *Anonymization*

Data anonymization is a technique that will not take away the original field layout (position, size and data type) of the data being anonymized, so the data will still look realistic in test data environments.

Anonymization has 3 primary goals[3]:

- To protect identities of specific user from being leaked
- To protect identities internal user from being revealed

- To protect specific security practices of organizations from being revealed.

Authentication	Anonymization	Reclamation

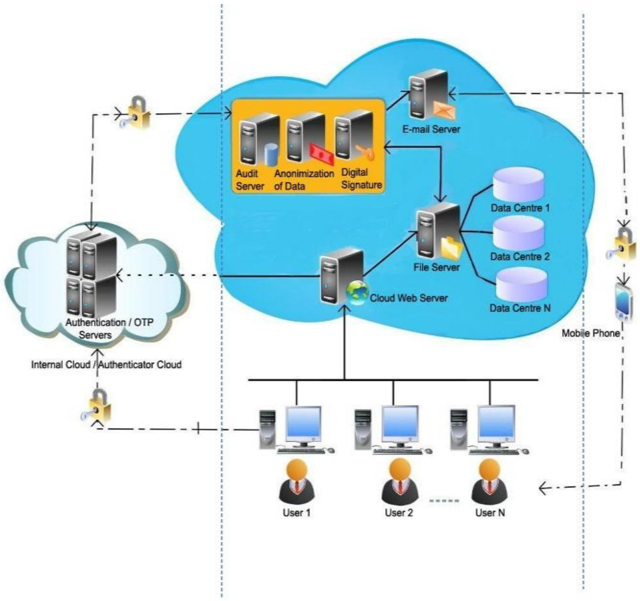


Fig 1: Architecture to Enhance Security of Data.

C. *Comparison Of Privacy Preserving Methods In Cloud Computing*

Several methods have been put forward to tackle the issue of privacy preserving. It is important, that the privacy has to be preserved anytime and anywhere[4]. Privacy preserving has originated as an important concerns with the reference to success of cloud computing. Privacy preserving deals with protecting the privacy of individual data or sensitive knowledge without sacrificing the utility of the data. Our paper provides multi attributes anonymization privacy preserving method used in cloud computing services.

Techniques	Description	crytographic techniques used
Anonymity-based method	Anonymises the sensitive data before storing in cloud	NO
Architecture for privacy-preserving database storage	Prevents both internal and external attacks	YES
Privacy-preserved Access Control	Determines access rights for users and achieves access control	YES
Privacy-preserving Authorisation System	Puts forth a policy based authorisation infrastructure	NO
Privacy Preserving Data Outsourcing	Guarantees privacy by means of data fragmentation	NO
PccP (Preserving cloud computing Privacy) approach	Preserves both user identification and information	NO
Dynamic reconstruction of metadata	Designs schemas for the database, Performs segregation and reconstruction of metadata	YES

Table:1.Privacy preserving in cloud

D. *Anonymization Techniques*

There are several techniques available to anonymize the data such as Encryption, substitution, shuffling, number and date variance and nulling some fields. We have discussed some anonymization techniques to obscure data in database[5].

1). *Data Hiding*

It suppresses a data value by replacing it with a value ‘0’. It is also called as Black marker anonymization. For example, while considering hospital database, an age of a patient may not be required for processing, so it is replaced with constant ‘0’.

2). *Hash Calculation*

It finds a hash value of either one field or several fields. It takes a variable input and produces fixed size hash of input. The MD5 or SHA can be used. For example, hash of first name and last name can be calculated.

3). *Shifting*

Shifting shifts a field or data value by specific value. It adds some offset to data value. Shift value is the only key to shift function, so that is kept secret. For example, an offset value 10 is added in age field.

4). *Data Enumeration*

Enumeration is also a substitution technique. It retains the chronological order in which events takes place. It is useful for applications demanding strict sequencing order. For example, salary field is enumerated while maintaining the order of execution.

5). *Ip Prefix-Preserving*

This method preserves the n-bit prefix on IP-address. Two anonymized IP addresses match on prefix of n-bits, if and only if two real IP addresses match on prefix of n-bits. The IP address is prefix preserved here.

Prefix-preserving anonymization belongs to Typed Transformation, which uses single anonymized value for each unique value of original data. The tool TCPdPriv uses prefix preservation anonymization.

we surveyed few anonymization methods to protect sensitive data in cloud. Formal models of security for anonymization are also discussed. Anonymization is a viable technique to secure cloud computing. It limits the misuse of sensitive data, but is not a complete solution to preserveconfidentiality. Research for anonymization and deanonymization is in process. The techniques which are currently safe for anonymization may fail in future. In future, the privacypreserving in cloud needs many efforts.

E. *Spectral Anonymization*

No existing anonymization algorithm provides both perfect privacy protection and perfect analytic utility.A spectral basis derived from the data's eigen vectors is one that can also provide substantial improvement. We introduce the term spectral anonymization to refer to an algorithm that uses a spectral basis for anonymization.We also propose new measures of privacy protection that are more general and more informative than existing measures, and principled reference standard with which to define adequate privacy protection.

III.EXISTING SYSTEM

The security level in existing system is not up to the mark. As we do not have any back up, consider the case of Orphanage project data reliability is not assured. As we store our data in database there are several issues in security. Some problems are lack of security, low data retrieval, data redundancy and consistency and no backup and recovery. For example attackers can easily steal the important details and financial reports of orphan home and sell it to other third party. As there is no backup of orphans data if data lost the orphan home center will face a bad impact where they may lose their important information for future analysis.

IV.PROPOSED METHOD

In the proposed system instead of storing our data in database we are storing in cloud to enhance space and security.our proposed system contains detail about orphans and their sponsors such as address location,phone number and contact details.so information stored in our system is very sensitive. Here we used anonymization to secure our sensitive data .Anonymization concept is used to enhance security in public cloud. This is an emerging technology which will bring about innovations in terms of business models and applications.Anonymity based method is used here it anonymizes the sensitive data before storing in cloud. This system provides safe and secure storing. We propose a novel methodology to schedule the data oriented grid applications in an efficient way to reduce the overall execution time of the applications[6]. Our proposed methodology consists of the following steps:

A.*Preprocessing*

In this step we read the semantic concepts and the meta data about the data sets available in the data grids. we remove the noisy semantics and meta data and clean the meta data to ensure the quality of semantic indexing.

B. *Grid Classification*

At this step with the output of preprocessing, we compute the similarity measure between the semantic concepts and the meta data of the data sets. Based on the similarity values we group the grids into set, which will be further used to look up the grids according to the requirement of the processes. We compute Euclidean distance to compute the similarity measure. The classified results are indexed into the data base. We use semantic indexing to store the classification results. In semantic indexing , the meta data are stored under a semantic concept based on the similarity value of meta data with the semantic concept.

C. *Algorithm for Grid Classification*

Step1: Read semantic concepts

Step2: Read meta data of data sets.

Step3: clean concepts and meta data.

Step4: compute similarity measure between semantic concepts and meta data.

Step5: Identify similar concepts and meta data.

Step6: Group similar meta data, under a related concepts.

D. *Scheduling*

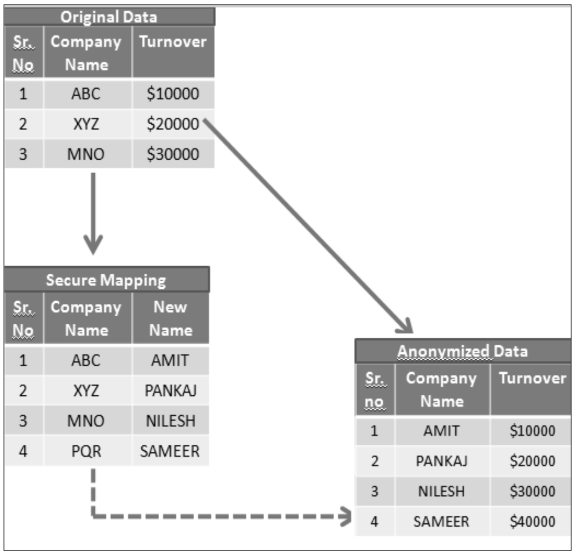
This step performs the scheduling of the process to help the process to be executed with in short time. The scheduler accepts the input job and retrieves the location of the data from the indexed data. Because of we use semantic indexing, its very easier for the scheduler to retrieve the location of the data objects. This reduce the scheduling time of the processes and also execution time will be less.

E. *Algorithm for Scheduling*

Step1: Read Input Job
Step2: Identify set of data objects necessary to execute the job.
Step3: compute similarity measure of data objects with semantic concepts.
Step4: Identify the semantic concept with respect to similarity measure.
Step5: retrieve the location of datasets from the indexed results.
Step6: return the results.

At the end of the scheduling process the application will be returned with the location of the grid where the application has to be executed. The query processor will post the process to the returned location and will wait for the result and return the result to the application.

F. *Data Anonymization in Cloud*



Original Data

Sr. No	Company Name	Turnover
1	ABC	$10000
2	XYZ	$20000
3	MNO	$30000

Secure Mapping

Sr. No	Company Name	New Name
1	ABC	AMIT
2	XYZ	PANKAJ
3	MNO	NILESH
4	PQR	SAMEER

Anonymized Data

Sr. no	Company Name	Turnover
1	AMIT	$10000
2	PANKAJ	$20000
3	NILESH	$30000
4	SAMEER	$40000

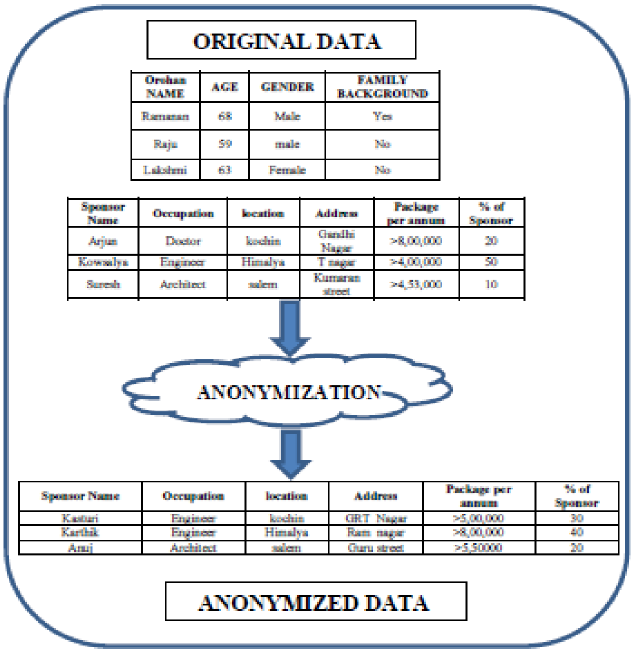
Fig 2: Data Anonymization

V. STORING THROUGH MOBILE PHONES

Since we are using the cloud computing technology, many devices can be grouped in the cloud including PC, Mac or smartphone. When the internet connection is available, it can send the data to the main server. If the internet connection is available directly it can send the candidate information directly to the server. Thus in our system we can store our data to the cloud at any place and also through other devices.

VII. RESULTS AND DISCUSSION

The final results shows that our proposed scheduling algorithm reduces the overall execution time of the application by reducing the scheduling time and execution time. Our indexing scheme reduces the scheduling time.



ORIGINAL DATA

Orohan NAME	AGE	GENDER	FAMILY BACKGROUND
Ramanan	68	Male	Yes
Raju	59	male	No
Lakshmi	63	Female	No

Sponsor Name	Occupation	location	Address	Package per annum	% of Sponsor
Arjun	Doctor	kochin	Gandhi Nagar	>8,00,000	20
Kowsalya	Engineer	Himalya	T nagar	>4,00,000	50
Suresh	Architect	salem	Kumaran street	>4,53,000	10

Sponsor Name	Occupation	location	Address	Package per annum	% of Sponsor
Kasturi	Engineer	kochin	GRT Nagar	>5,00,000	30
Karthik	Engineer	Himalya	Ram nagar	>8,00,000	40
Anuj	Architect	salem	Guru street	>5,50000	20

ANONYMIZED DATA

Fig 3: Result and discussion

CONCLUSION

This paper identifies the importance of privacy in cloud computing and provides privacy preserving technologies used in cloud computing. Privacy should be built into every stage so that it can be implemented in cloud environment for orphanage management system. Lots of techniques for anonymization have been implemented, but still there is a fear of security breach. So we introduced Spectral anonymization here to provide perfect privacy protection and analytic utility. The great challenge for an anonymization scheme is to provide adequate privacy protection while minimally affecting the data's analytic utility.The concept of spectral anonymization is therefore attractive due to its simplicity and power.The proposed method achieves good results and reduces the overall execution time. We further analyze the query execution process to reduce the over all execution time.

REFERENCES

[1] B. C. M. Fung, K. Wang, P .S. Yu , R.Chen, Privacy Preserving data publishing : A survey of Recent Developments,"ACM Comput.Survey,vol 42,No 4,PP.I53,2010.

[2]. Jiajin Le, Shuo Jiang, Yan Zhao and Jian Wang, Providing Privacy Preserving in Cloud Computing, Proc.of IEEE Infocom, pp 472-475, 2010.

[3].R.Pang , M.Allman, V.Paxson, and Lee, “The Devil And Packet Trace Anonymization”, ACM Computer Communication Review, 36(1):29–38, January 2006.

[4].T,Jothi leela and N,Saravanan, “Ptivacy Preserving Approaches in Cloud: A survey, IJST, vol.6.

[5].V.K.Saxena, Dr.Shashank Pushkar, ”Anonymization Approach for privacy preserving in cloud computing”, International Conference on cloud, Big Data And Trust 2013. Nov 13-15.

[6].L.Gomathi , M.Maheswari. "A Semantic Based Scheduling Algorithm for Data Intensive applications on Global cloud". International Journal of Engineering Trends and Technology (IJETT). V4(8):3527-3530 Jul 2013.

[7].S.Ramgovind, E. Smith,” The Management of Security in Cloud Computing”,IEEE Internationl Conference on Cloud Computing, 2010.