International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 7 26 - 31

Running Big Data Privacy Preservation in the Hybrid Cloud Platform

Onkar S. Undale
Department of Computer Engineering
Dhole Patil College of Engineering
Pune, India
Onkar.undale@hotmail.com

Prof. Bharati Kale
Department of Computer Engineering
Dhole Patil College of Engineering
Pune, India
Bharatikale02@gmail.com

***Abstract*—** Now a day's cloud computing has been used all over the industry, due to rapid growth in information technology and mobile device technology. It is more important task, user's data privacy preservation in the cloud environment. Big data platform is collection of sensitive and non-sensitive data. To provide solution of big data security in the cloud environment, organization comes with hybrid cloud approach. There are many small scale industries arising and making business with other organization. Any organization data owner or customers never want to scan or expose their private data by the cloud service provider. To improve security performance, cloud uses data encryption technique on original data in public cloud. Proposed system work is carried out how to improve image data privacy preserving in hybrid cloud. For that we are implementing image encryption algorithm based on Rubik's cube principle improves the image cryptography for the public cloud data security
.

Keywords- Big data; Image data; Cloud service provider; Hybrid cloud; Privacy preserving; Rubik's cube principle

*****

I. INTRODUCTION

Now a day's people are living with many Internet organization such as Google, Amazon, Facebook, using mobile devices and sensors. They are connected to each other, data is generated, accessed and shared with each other. Many sectors are working with cloud technology. It shows that cloud computing can provide a great business model in term of cost effective resolutions feature with big data. Providing the benefits of cut down resource utilization cost through sharing of resources for the computational and storage purpose, on demand resource provisioning can be possible, pay only for what you are using. Big data is collection of massive amount of relevant and non-relevant data. Processing of big data, extracting information from collection of data violates the policy of data privacy preservation. So, a question is what type of security and privacy preserving technology is adequate suitable for efficient direct access to big data [1]. As we consider with respect to privacy preservation of data stored in the public cloud environment slow down the adoption of cloud environment for the big data. However, the social network or medical system carries sensitive information. These are the issues are concerned with big data volume, velocity, variety of data and cloud infrastructure to be utilize, migration of data in the cloud from different data source and format demotivate the use of traditional security mechanism. AES is a standard cryptographic algorithm which is most commonly used encrypt the user's data and store it to public cloud [3]. But when we are considering image data, it has larger size as compare to text data. If file size larger then more computational power is required. Hence it demotivates the mobile devices to this approach even though it is standard approach, because it leads much more battery power consumption due to heavy workload. Automatically it will increase the delay due to limitation of computational power. Hence we can say that traditional cryptographic mechanisms are degrade the overall performance of system [4]. New concept of hybrid cloud infrastructure objective is to go and protect data, when data is process and becomes information, so data is to be protected and should not be corrupted. Data manipulation is done at the cloud application level. There are many image cryptographic algorithm discovered in recent years to improve speed of process, among these chaos-based mechanism, substitution-diffusion layout is most suitable mechanism [5]. In the step of substitution, pixel positions are shifted with respect to some chaotic map. In the next step of diffusion change the pixel value from shuffled image through chaotic sequence. Encryption of image can be done with one to one mapping function of Rubik's cube principle performance of system is improved (i.e. It come up with new technique of confusion and diffusion property to overcome the traditional substitution and diffusion) Requirement of random parameters for mapping the function are stored at private cloud.

II. LITERATURE REVIEW

A paper created by Qingchen Zhang, Laurence T. Yang and Zhikui Chen on Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning. In this paper, to provide security to the private data BVG encryption mechanism is used for the encryption of data. This architecture hires the cloud server to execute computational operation on algorithmic encrypted data. sigmoid function is a polynomial run time function which performs secure computational operation on BVG encryption. This system is based on public cloud environment [2]. Encryption and decryption operation performed at client's mobile devices and rest of computational tasks is performed at the cloud environment. In the next paper of Shared Authority Based Privacy-Preserving Authentication Protocol in Cloud Computing. survey conducted by Hong Liu, Huansheng Ning, Qingxu Xiong, Laurence T. Yang focused on how to provide authentication process for cloud server to access by requested users sharing data [2]. Accessing permission mechanism is most important factor for this privacy preservation system. This can be achieved by setting up certain rules and procedures author called it shared

authority based privacy preserving authentication protocol (SAPA) [3]. Author Hui Zhu, Rongxing Lu, Cheng Huang, Le Chen and Hui Li proposed a system of an Efficient Privacy-Preserving Location Based Services Query Scheme in Outsourced Cloud [4]. In this paper they provide the mechanism in the form of outsource the data into cloud for computation. This approach is most convenient for people who love to work with mobile devices. In this mechanism with help of mobile device location based query is outsource to the cloud is used. Author named it as EPQ. i.e. location based service (LBS) helps to outsource the EPQ query data to the cloud server first time in encrypted format. Then later only registered users are able to get the LBS query response without breaking of his or her location information in between LBS provider and the server of cloud [5]. To improve homomorphic encryption, it uses special query algorithm (SRQC) operate over the cipher text work with EPQ and LBS. Author ensures EPQ is able to provide privacy preservation for users query submission secrecy of data at LBS while outsourcing to the cloud server.

III. PROBLEM STATEMENT

In term of the privacy preservation, data stored in public cloud slow down the adoption of cloud for big data, because of sensitive information [1]. The data owner themselves do not want any other people to scan their data or expose it. Since data volume is huge and mobile device are widely used. The traditional cartographic approach is not suitable for big data time consuming heavy workload leads to extra battery power consumption for processing information [1][2].

IV. IMPLEMENTATION DETAILS

A. *Existing System Architecture*

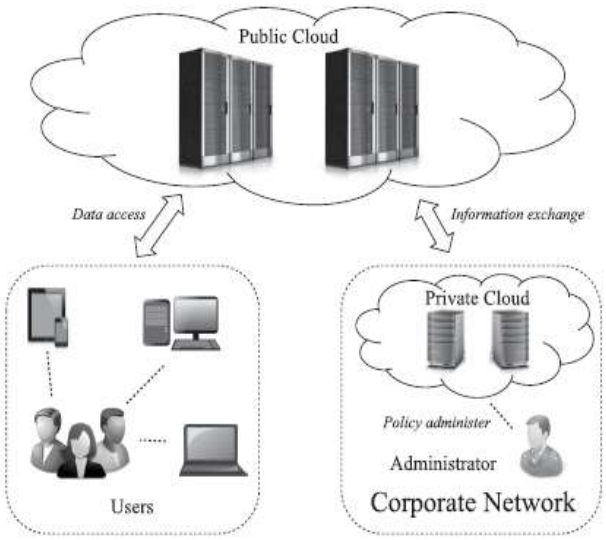

Figure 1. Existing System Architecture

Existing system provides hybrid cloud environment and RBE scheme role based encryption. As shown in figure 1 existing system, public cloud has more priority than organizational private cloud [4]. In this system role hierarchy and user's membership information are considered as sensitive data. which is stored at private cloud and user's actual confidential data stores in encrypted format at public cloud as non-sensitive data. All end users connected first to the public cloud to access information and then RBE scheme is applied [7]. Hence we can say that there is attacker and CSP threat to this system, because entire user's data stored at public cloud

B. *Proposed System Architecture*

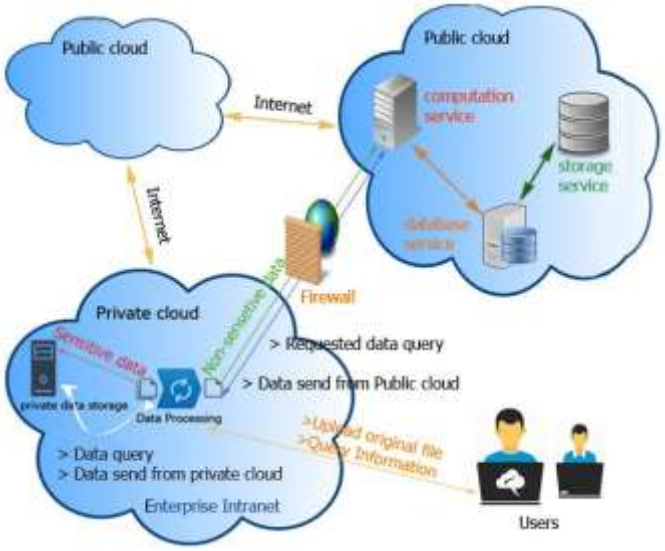

Figure.2 Proposed System Architecture

This privacy preservation mechanism is based on existing cloud computing architecture. There are two components for this architecture namely cloud service provider (CSP) and cloud customer [8]. As shown in figure 2 data generated at private cloud servers, user have to analyses the sensitive and non-sensitive data. If data is non sensitive it will forward to the public cloud and sensitive data will be stored at private cloud [9]. At the stage of data retrieval for users query response both public and private cloud will communicate to generate the result.

V. DESIGN GOAL

A. *Problem Description*

use of hybrid cloud environment for image data privacy preservation. So, by differentiating sensitivity and non-sensitivity of data form original data. store them separately in trusted private or public cloud with respect to location, after applying Rubik's cube principle based image encryption algorithm. Some performance factor we have to consider for hybrid cloud and reduce the following overheads: (1) load ratio of data stored in private cloud, (2) private public cloud communication overhead. (3) Over all delay required to complete data life cycle [3][5][6].

B. *Mathematical Model*

Consider a set of users U registering with the cloud. These users can perform tasks such as file uploading, downloading. The set U can be represented as follows,

$$U = \{u1, u2, u3, ...un\}$$

A user can upload multiple image files on the cloud, so here we can represent a set of files F as,

$$F = \{f1, f2, f3, ...fn\}$$

When a user uploads a file to cloud, at cloud the file gets divided into multiple parts. Here we are dividing file into 4 parts. (i.e. each part size will be)
File size / Required bits in each block =4 Consider a set of threads T represented as

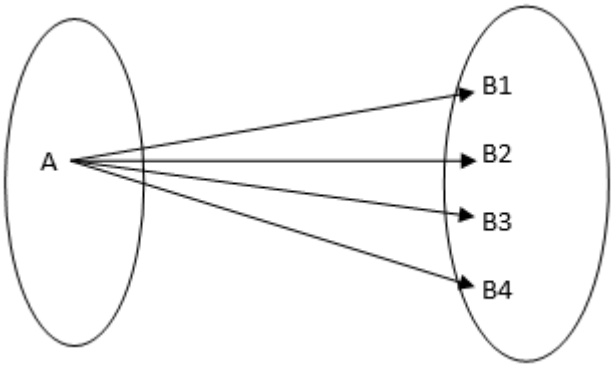

Figure. 3. One-to Many Function

Consider a set of threads T represented as follows performing operation on each block of image. T= {t1, t2, t3...tn} Here we are computing summations of all row pixel and applying module 2 operation on that result we are rotating pixel left shift or right shift in 1-dimensional array. Then using XOR operation for second time shuffling pixel for rows with another randomly generated encryption array [10]. Same steps of implemented for all four blocks and finally we are combining result in to encrypted image file [11]. For decryption, we have to perform reverse operation on same file. The proposed architecture workflow is as shown in given state transition diagram. Here in Figure 4. There are five states are there,
Q0: Private cloud
Q1: Image data encryption state
Q2: Image data decryption state
Q3: Private cloud parameter storage
Q4: Public cloud storage
Q5: Public cloud

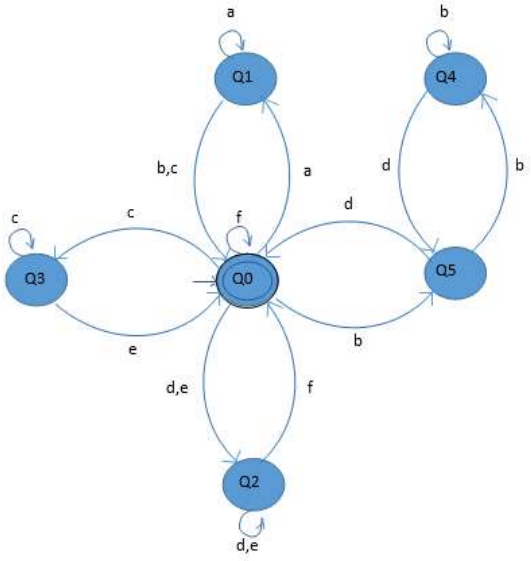

Figure 4. NFA State Diagram for Mathematical Workflow of Architecture.

Table I: Workflow transition table for each state

String i/p	Q \ Σ	a	b	c	d	e	f
a*f*	q0	q1	q4	q5	q2	q2	q0
bbbb*ccc*	q1	q1	q0	q0	qϕ	qϕ	qϕ
d*e*ff*	q2	qϕ	qϕ	qϕ	q2	q2	q0
b*dddd*	q3	qϕ	q3	qϕ	q4	qϕ	qϕ
bb*dd*	q4	qϕ	q3	qϕ	q0	qϕ	qϕ
c*eee*	q5	qϕ	qϕ	q5	qϕ	q0	qϕ

The above transition diagram and table explains how workflow is carried out in our proposed system. State Q0 is an initial and final state of workflow. It is present at private cloud environment. State Q5 is the state of public cloud. Both states are communicating to response the users query submission. Therefor we can say that these two states are in heavy communication workload. User are connected at private cloud. Other states in workflow connected to Q0 such as Q1 is an encryption state. Q2 is state for to store the encryption parameter or key value of cipher text data.
(sensitive and non-sensitive data). State Q3 is decryption state. Last state Q4 which is connected to Q5 state of public cloud is used to store the non-sensitive data cipher text data at public cloud storage.

VI. ALGORITHM

A. Rubik's Cube Principle Based Image Encryption

In the privacy preservation sometimes image data is in the confidential nature, user don't want to expose it. This idea has been growing around the use of Rubik's cube principle image cryptography. This algorithm is very useful for digital communication system mobile devices, image sensors [12]. We are using these devices for encryption of sensitive image data [13]. This algorithm provides quality of secure encryption mechanism with minimal computational power on mobile devices. the work carried out of this algorithm is as follows:

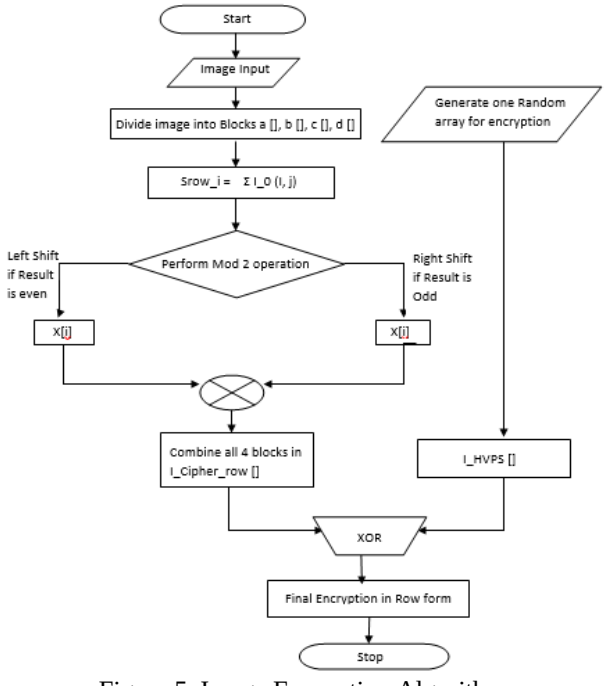

Figure 5. Image Encryption Algorithm

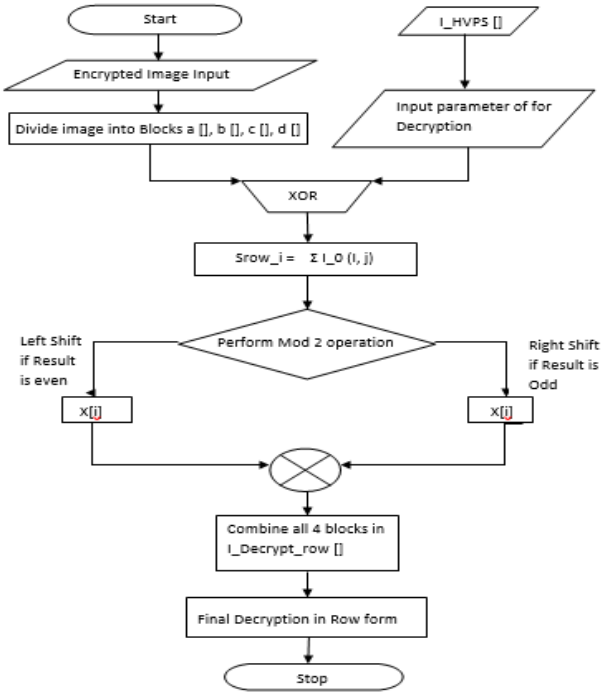

Figure 6. Image Decryption Algorithm

- Image will be converted to one-dimensional array.
- The array is divided in to 4 parts, Then Sum of each block will be calculated and modulo by 2
- If the modulo of current block/part is 0 then it is right circular shifted else left circular shifted.
- Then we generate secure random byte array having size of part.
- The secure random byte array is XORed with all parts which will be our encrypted image parts.
- Parts are stored in reverse order (i.e. 4,3,2,1)
- This will generate a scrambled image.
- Finally, while recovering the image the same algorithm is used the change is parts are recorded into their original order (i.e. 1,2,3,4)

B. Complexity Analysis

We have measured encryption time for different sized gray scale and color images by using the proposed image encryption algorithm. Computational analysis has been done on a 2.40 GHz Intel Core TM i5-3630QM Dell Inspiron laptop.

Table II
Running time for proposed image encryption

Image size	Dimensions	Image encryption time	Image Decryption Time
858 kb	1024×768	600 MS	115 MS
826 kb	1024×768	590 MS	90 MS
267 kb	512×360	20 MS	31 MS
132 kb	238×126	7 MS	0.32 MS

From Table II we can say that results are provided in previous reference AES image encryption algorithm runs slower than our system. [14] The Rubik's cube principle based image encryption algorithm needs O(n2) time complexity [15]. We have proved this proved this theoretical complexity in substitutions and implementing this algorithm as shown in Table II and Figure 7 and 8. In this algorithm we required O(n) time complexity to read the RGB value of image, time complexity for counting, making proxy image array generation is also a O(n) [16]. Other image operation such as merge, rotate pixel, XOR all these mathematical operations are done in constantan time.

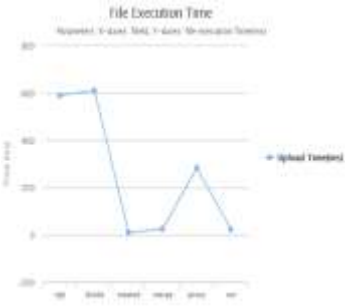

Figure 7. 858 kb file size image encryption time

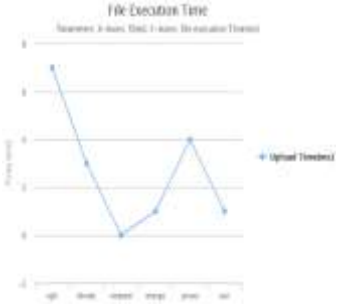

Figure 8. 132kb file size image encryption time

As shown above Figure 7 and 8 shows require time for image encryption with different image file as shown in Table II. Since we can say that our image encryption algorithm takes polynomial time complexity.

VII. EXPERIMENTAL RESULTS

In this section, we are expressing the result that we are carried out to test the efficiency of image encryption with respect to its security. Image visual quality test and Analyzing encrypted image security test these two test are required to express the proposed system result.

A. Image visual quality test

To conduct this test, we taking images with different sizes and their quality. We observe that after applying our algorithm pixel positions of images are encrypted pixel by pixel which will give quality of encryption as shown in fig 7 and fig 8.

Figure 9. Original Image

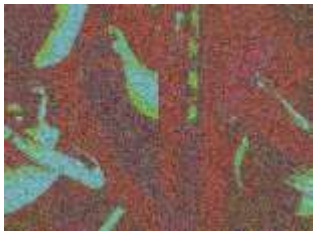

Figure 10: Shuffling Pixel

Figure 11. Fully Encrypted Image After XORing Operation

Figure 12. Final Decrypted Image

As shown in Table II and Figure 9 with file size 858kb, 1024×768-dimension image is taken for encryption. We see that performance of image encryption is very good as compare to another image cryptographic algorithm.

B. Analyzing encrypted image security test

Considering security parameter for image cryptography there are many methods are present to break image encryption, brute force attack, pain text or cipher text attack. Mainly encryption is dependent on two factors one in key space and another is key sensibility or dependability. In key space factor provides no of possibility key combination for brute force attack. It is impossible to make a brute force attack on encrypted image with our proposed algorithm because it requires large no of key space to decrypt it. Second factor key sensibility here good encryption always provides high feature of key sensibility, i.e. small pixel level changes made to large difference show in encrypted image as shown in Figure. 16 and for this we are generated same size key for encryption form given original image. So, we can say that our algorithm is highly key sensible without key it won't be work for decryption process.

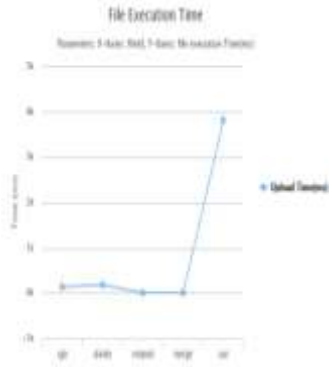

Figure 13. Upload Time

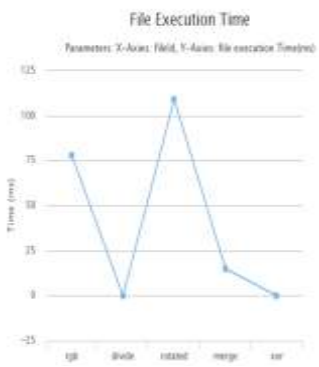

Figure 14. Download Time

We are performing each pixel level mathematical operation which changes the original image pixel value XORing with proxy parameter array pixel value and produce a encrypted image. In case original color RGB value representation in hexadecimal representaion value get coverd to perform XOR operation on each pixel. The same numberd binary addtion of XOR opration value gives 0 resultant pixel value representation.

Figure 15. Original Image

Figure 16: Shuffling Pixel

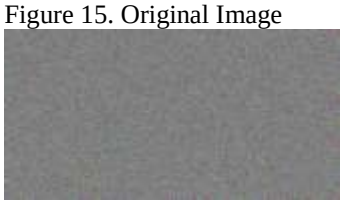

Figure 17. Fully Encrypted Image.

Figure 18. Final Decrypted Image

In above most of example images have good color pixel which helps to recover the encrypted pixel. so, we can say that this algorithm supports good color saturation pixel cryptography.

VIII. CONCLUSION

As more applications are migrating into cloud, it is imperative to migrate big data workflow into cloud to take advantage of cloud scalability and also to handle the ever increasing data scale and analysis complexity of such applications. Hybrid cloud offers unprecedented scalability to big data privacy preservation workflow system and cloud potentially change the way we perceive and conduct the image data encryption algorithm based on Rubik's cube principle. The scale and complexity of science problem that can be handled greatly increased on the cloud.

We proposed our implementation image data privacy preservation on Amazon EC2 cloud platform and also present our efforts in reducing delay of communication overhead between public and private cloud. In which client side cloud environment service management, monitor and developed.

ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for their valuable suggestions that significantly improved the quality of this paper. I also thank Prof. Bharati Kale who provided me the opportunity to present this Dissertation.

REFERENCES

[1] Xuyun Zhang, Chang Liu, Surya Nepal, Suraj Pandey, and Jinjun Chen, Member, IEEE, A Privacy Leakage Upper Bound Constraint-Based Approach for Cost-Effective Privacy Preserving of Intermediate Data Sets in Cloud, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 24, NO. 6, JUNE 2013, 1045-9219/13/$31.00 2013 IEEE Published by the IEEE Computer Society

[2] Lan Zhou, Vijay Varadharajan, and Michael Hitchens, Achieving Secure Role-Based Access Control on Encrypted Data in Cloud Storage, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 8, NO. 12, DECEMBER 2013, 1556-6013 © 2013 IEEE

[3] Piotr K. Tysowski and M. Anwarul Hasan, Senior Member, IEEE, Hybrid Attribute- and Re-Encryption-Based Key Management for Secure and Scalable Mobile Applications in Clouds, IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 1, NO. 2, JULY-DECEMBER 2013, 2168-7161/13/$31.00 2013 IEEE Published by the IEEE CS, ComSoc, PES, CES, & SEN

[4] Hong Liu, Student Member, IEEE, Huansheng Ning, Senior Member, IEEE, Qingxu Xiong, Member, IEEE, and Laurence T. Yang, Member, IEEE, Shared Authority Based Privacy-Preserving Authentication Protocol in Cloud Computing, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 26, NO. 1, JANUARY 2015, 1045-9219 2014 IEEE

[5] Mark-John Burke, Anne V.D.M. Kayem, K-Anonymity for Privacy Preserving Crime Data Publishing in Resource Constrained Environments, 2014 28th International Conference on Advanced Information Networking and Applications Workshops, 978-1-4799-2652-7/14 $31.00 © 2014 IEEE DOI 10.1109/WAINA.2014.131

[6] Yong Zhao, Member, IEEE, Youfu Li, Student Member, IEEE, Ioan Raicu, Member, IEEE, Shiyong Lu, Senior

Member, IEEE, Cui Lin, Member, IEEE, Yanzhe Zhang, Wenhong Tian, Member, IEEE, and Ruini Xue, Member, IEEE, A Service Framework for Scientific Workflow Management in the Cloud, IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 8, NO. 6, NOVEMBER/DECEMBER 2015, 1939-1374 2014 IEEE.

[7] Mohamed Hefeeda, Senior Member, IEEE, Tarek ElGamal , Kiana Calagari, and Ahmed Abdelsadek, Cloud-Based Multimedia Content Protection System, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 3, MARCH 2015, 1520-9210 © 2015 IEEE.

[8] Qingchen Zhang, Laurence T. Yang and Zhikui Chen, Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning, IEEE TRANSACTIONS ON COMPUTERS, FEB 2015, 0018-9340 (c) 2015 IEEE.

[9] Zhihua Xia, Member, IEEE, Yi Zhu, Xingming Sun, Senior Member, IEEE, Zhan Qin, Member, IEEE and Kui Ren, Senior Member, IEEE, Towards Privacy-preserving Content-based Image Retrieval in Cloud Computing, IEEE TRANSACTIONS ON COMPUTER COMPUTING, SEPTEMBER 2015, 2168-7161 (c) 2015 IEEE.

[10] Valeriu Manuel IONESCU, Adrian-Viorel DIACONU, Rubik's cube principle based image encryption algorithm implementation on mobile devices, ECAI 2015 - International Conference – 7th Edition Electronics, Computers and Artificial Intelligence 25 June -27 June, 2015, Bucharest, ROMÂNIA, 978-1-4673-6647-2/15/$31.00 ©2015 IEEE.

[11] Hui Zhu, Member, IEEE, Rongxing Lu, Senior Member, IEEE, Cheng Huang, Le Chen, and Hui Li Member, IEEE, An Efficient Privacy-Preserving Location Based Services Query Scheme in Outsourced Cloud, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, 0018-9545 (c) 2015 IEEE.

[12] Chao-Yung Hsu, Chun-Shien Lu, and Soo-Chang Pei, Fellow, IEEE, Image Feature Extraction in Encrypted Domain with Privacy-Preserving SIFT, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 11, NOVEMBER 2012, 1057–7149/$31.00 © 2012 IEEE.

[13] Amit Mukherjee, Miguel Velez-Reyes, Senior Member, IEEE, and Badrinath Roysam, Senior Member, IEEE, Interest Points for Hyperspectral Image Data, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 47, NO. 3, MARCH 2009, 0196-2892/$25.00 © 2009 IEEE.

[14] Ahmed B. Mahmood, Robert D. Dony, Segmentation Based Encryption Method for Medical Images, 6th International Conference on Internet Technology and Secured Transactions, 11-14 December 2011, Abu Dhabi, United Arab Emirates, 978-1-908320-00-1/11/$26.00 ©2011 IEEE.

[15] Yuriy Brun, Member, IEEE, and Nenad Medvidovic, Member, IEEE, Entrusting Private Computation and Data to Untrusted Networks, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 10, NO. 4, JULY/AUGUST 2013, 1545-5971/13/$31.00 2013 IEEE.

[16] WENJUN LU1, AVINASH L. VARNA2, (Member, IEEE), AND MIN WU3, (Fellow, IEEE), Confidentiality-Preserving Image Search: A Comparative Study Between Homomorphic Encryption and Distance-Preserving Randomization, Received December 15, 2013, accepted January 15, 2014, date of publication February 20, 2014, date of current version March 4, 2014. Digital Object Identifier 10.1109/ACCESS.2014.2307057, 2169-3536 2014 IEEE.

[17] Michael Menzel, Member, IEEE, Rajiv Ranjan, Member, IEEE, Lizhe Wang, Senior Member, IEEE, Samee U. Khan, Senior Member, IEEE, and Jinjun Chen, Senior Member, IEEE, CloudGenius: A Hybrid Decision Support Method for Automating the Migration of Web Application Clusters to Public Clouds, IEEE TRANSACTIONS ON COMPUTERS, VOL. 64, NO. 5, MAY 2015, 0018-9340 2014 IEEE.

[18] Xuyun Zhang, Wanchun Dou, Jian Pei, Fellow, IEEE, Surya Nepal, Member, IEEE, Chi Yang, Chang Liu, and Jinjun Chen, Senior Member, IEEE, Proximity-Aware Local-Recoding Anonymization with MapReduce for Scalable Big Data Privacy Preservation in Cloud, IEEE TRANSACTIONS ON COMPUTERS, VOL. 64, NO. 8, AUGUST 2015, 0018-9340 2014 IEEE.

[19] Adrian-Viorel Diaconu and Khaled Loukhaoukha, An Improved Secure Image Encryption Algorithm Based on Rubik's Cube Principle and Digital Chaotic Cipher, hindawi journals mpe.2013,848392.

[20] M.Sirisha, SVVS Lakshmi, Pixel Transformation based on Rubiks Cube Principle (IJAIEM), Volume 3, Issue 5, May 2014.

[21] Xueli Huang and Xiaojiang Du, Achieving Big Data Privacy via Hybrid Cloud, 2014 IEEE INFOCOM Workshop on Security and Privacy in Big Data 978-1-4799-3088-3/14

[22] Bayardo RJ, Agrawal R (2005) Data privacy through optimal k-anonymization. In: Proceedings of the 21st IEEE international conference on data, engineering (ICDE05), pp 217228

[23] Samiksha Shuklat, G Sadashivappa, A Distributed Randomization Framework for Privacy Preservation in Big Data, 978-1-4799-3064-7/14/$31. 00©20 14 IEEE