ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

Privacy-Preserving Data Encryption Strategy for Big Data in Mobile Cloud Computing Environment

Sumit Vikram Tripathi[1], Ritukar[2], Prof. Murthy B[3], Dr. K. S. Jagadeesh Gowda[4]

U.G. Student, Department of Computer Science & Engineering,Sri Krishna Institute of Technology, Bengaluru, India[1]

U.G. Student, Department of Computer Science & Engineering,Sri Krishna Institute of Technology, Bengaluru, India[2]

Assistant Professor, Department of Computer & Engineering,Sri Krishna Institute of Technology, Bengaluru, India[3]

Professor, Department of Computer & Engineering,Sri Krishna Institute of Technology, Bengaluru, India[4]

ABSTRACT:Technology is enhancing each and every day especially in the field of Information Technology and data is very momentous elements. The large volume of data generated through devices is a major obstacle to handle in real time. The accomplishment of data ciphering is a crucial problem during the data progression and dissemination. In order to achieve pursuance, many application disregards data encryption. This paper represents a concern about data privacy and suggests a novel data encryption approach known as Dynamic Data Encryption Strategy (DDES). This accession is drafted to magnify privacy protection scope within liquidate time constraints.

KEYWORDS:Privacy-preserving, data encryption strategy, big data, mobile cloud computing, cyber security

I. INTRODUCTION

With the evolution of the Internet, the world ensures that we are all connected at the touch of the button. Mobile Computing is future of the technology as it allows us to connect to the Internet and the data that demand distribution. Mobile cloud computing techniques legitimates diverse applications. Although mobile cloud computing has several advantages but a little worry about safeguard and confidentiality of data.

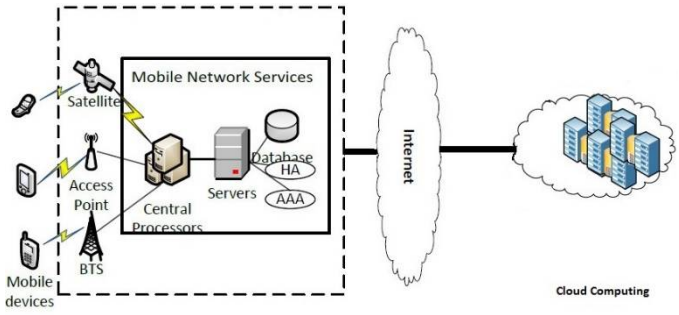

Fig. 1: Cloud Computing

One of the privacy concerns is caused by unencrypted data transmissions due to the large volume of data. This phenomenon can result in privacy leakage issues since plaintexts are unchallenging for adversaries to capture information in a variety of ways, such as jamming, monitoring, and spoofing. The proposed model is called Dynamic Data Encryption Strategy (DDES) model, which is designed to protect data owners' privacy at the highest level when using the applicable devices and networking facilities. Two crucial terms are designed for implementing the data encryption strategy, which includes Paired Data and Pairs Matching Collision. In addition, two crucial algorithms are

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

proposed for supporting the implementation of the DED algorithm, which is Weight Modelization (WM) Algorithm and S Table Generation (STG) Algorithm. These two new algorithms further identify the methods of identifying privacy values when making a determination on encrypting the input data. The main contributions of this work are threefold. 1) This work proposes a novel approach that selectively encrypts data packages to maximize the privacy protection level under timing constraints in big data. Two working modes are considered when creating the transmission strategy, including encryption and non-encryption modes. 2) The proposed algorithm offers an optimal solution providing the maximum value of total privacy weights. Two involved constraints are execution time and privacy levels. 3) The findings of this research provide big data-based solutions with an adaptive transmission approach focusing on protecting privacy.

II. RELATED WORK

Researches addressing the attacks in social networks have been paid attention by many scholars. First, Zhang et al. [1] proposed an approach named SCLPV for cloud-based Cyber Physical Social Systems (CPSS) to avoid malicious auditors. Second, Wang et al. [2] focused on developing an approach offering a secure cloud system that could support privacy preserving public auditing. Despite there are a few researches focusing on the impact of human behaviours in privacy issues, the unencrypted data still leave adversaries a great chance to steal users' privacy. Furthermore, privacy concerns can be caused by various dimensions in mobile clouds. Untrustworthy data is the first aspect of creating privacy leakages that can be hardly perceived by users or service providers due to two main reasons. The first reason is that it is difficult to identify the collected data because of the low trustworthy. The other one is that adversaries do not distribute any identification information such that it is hard to generate threat alerts. Next, the data analytic technique is considered a potential solution to identify trustworthy data while the data size becomes large, which has been addressed by the prior research. Nevertheless, these researches mainly concentrate on threat detections by using a variety of analysis techniques. Our research proposes an approach that aims to increase the rate of encrypted data while considering the value of the data encryptions. In addition, the vulnerability detection is also an important aspect of preventing privacy leakage [3]. Mulliner et al. [4] proposed a detective approach that focused on the vulnerabilities caused by the instances of Graphical User Interface (GUI) element misuse. This method considered the misuses of the GUI element attributes in the GUI-based application context. In the context of big data, an efficient privacy policy compliance checking mechanism is a significant part in building up a secure searching system [5]. Lack of tracking functionality in Web browsers can result in privacy issues since adversaries are not surveilled under most current operating environments. An efficient secure networking system can also reduce the rate of the threat amplifications [6].

However, there are a variety of vulnerabilities even though many access control models have been developed. Data transmissions in wireless networks create a large number of opportunities for attackers to intrude the communications and steal data. Privacy can be threatened even some data segments are captured by adversaries because of the advanced data mining techniques.

In this paper, we develop an approach that selectively encrypt data in order to protect privacy. The data encryptions depend on the return value of the encryptions and data attributes in order to minimize the chance of privacy leakage when adversaries apply data mining techniques.

III. CONCEPTS AND PROPOSED APPROACH

3.1 Problem Definition: To selectively encrypt maximum data packets under strict time constraints

Inputs:

1. *Data package types {Di}*
2. *The number of data for each data package type{NDi}*
3. *Execution time when encrypting data for each single data {T e Di}*
4. *Execution time without encryptions for each single data {T n Di}*
5. *The privacy weight value for each data type {WDi}*

Outputs: A strategy determining which data will be encrypted under a given timing constraint.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

Therefore, output is an encryption strategy that determines which data packages should be encrypted with in time constraints.

3.2 Dynamic Data Encryption Strategy Model

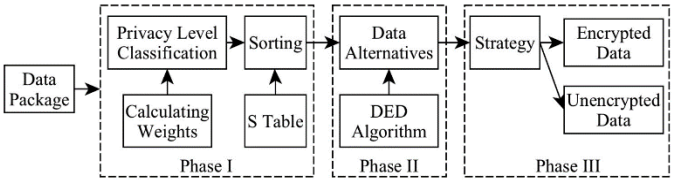

Fig. 2: Phases of Dynamic Data Encryption Strategy

There are mainly **three phases** forming the solution:
Phase I : Sorting by Weights
Phase II : Data Alternatives
Phase III: Output

3.2.1: Phase I: Sorting by Weights

This is the first phase of the proposed model. All the data package types are sorted at this phase. The sorting operations consider two very important attributes. Both execution time and privacy protections; therefor, two variables are involved, which are PWVs and the corresponding encryption execution time.

In order to improve the level of privacy protection, we introduce a mechanism called Pairs Matching Collision (PMC). This mechanism is designed to avoid the scenario when two plain texts can release users' privacy even though leaking each plain text will not be harmful. The operating principle of PMC mechanism is to make sure that the two pre-defined pair data have at least one data encrypted. The paired data must contain privacy information when they are transmitted or operated in plain texts.

3.2.2: Phase II: Data Alternatives

In this phase we select the data packages for encryption operations. We propose the DED algorithm to accomplish this phase. S Table will be used for providing the reference of protection efficiencies. There is a simpleoperating principle and that is that data package with higher value of SDi has a higherlevel alternative priority than those data packages having lower values of SDi . There are a few sub-steps for selecting data packages.

First, a timing scope needs to be identified.

Second, data alternatives are executed.

3.2.3: Phase III: Output

This phase producesasoutput an encryption strategy deriving from the correct outcomes of Phase II. Those data with higher-level encryption priority will be selected for the encryptions under a certain constraints. The rest of data will not be encrypted such that plain texts operations are applied.

IV. ALGORITHMS

The main algorithms used in our DDES model, which include Dynamic Encryption Determination (DED) algorithm, S Table Generation (STG) algorithm, and Weight Modelization (WM) algorithm.

DED algorithm is designed to dynamically select data packages that can be encrypted under certain conditions when considering both timing constraints and resources capacities. STG and WM algorithms are designed for supporting DED algorithm.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

4.1 Dynamic Encryption Determination Algorithm

This algorithm is used to dynamically select data packages that can be encrypted under certain conditions when considering both timing constraints and resources capacities

Require: *S Table*, *M-Table'*, T_c, T_m
Ensure: P (Encryption Strategy Plan)

Input *S Table*, *M Table*, T_c, T_m
Initialize $P \leftarrow \emptyset$
$T_s \leftarrow [T_c - (T_m + \sum_{D_i \in \text{S Table}} (N^n_{D_i} \times T^n_{D_i})$
$\quad + \sum_{D_i \in \{W_{D_i}=0\}} (N^n_{D_i} \times T^n_{D_i}))]$
/*In line with Eq. (5)*/
while *S Table* is not empty **do**
  Get D_i having the highest priority from *S Table*
  for $\forall$ D_i, i=1 to N_{D_i} **do**
    if $T_s > T^e_{D_i} - T^n_{D_i}$ **then**
      Add one D_i to P
      $T_s \leftarrow T_s - (T^e_{D_i} - T^n_{D_i})$
    else
      Break
    end if
  end for
end while
Output P

4.2 Weight Modelization Algorithm

This algorithm is designed to support DED algorithm.

Require: *M Table*, *Co-Table*
Ensure: *M-Table'*

Input *M Table*, *Co-Table*
for $\forall D_i$ in *M Table* **do**
  if D_i is in *Co-Table* **then**
    Get the pairs matching collisions ($D_i \leftrightarrow D_j$)
    if D_j is in *M Table* **then**
      if $T^e_{D_i} < T^e_{D_j}$ **then**
        $W^e_{D_i} = \infty$
      else
        $W^e_{D_j} = \infty$
      end if
    end if
  end if
end for
Output *M-Table'*

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

4.3 S Table Generation Algorithm

Thisalgorithms is designed to support DED algorithm.

Require: *M-Table'*
Ensure: *S Table*, T_m

Input *S Table*
Initialize *S Table* $\leftarrow \emptyset$
Initialize $T_m \leftarrow 0$
for $\forall D_i$ in *M-Table'* **do**
  if $W^e_{D_i} = \infty$ **then**
    $T_m \leftarrow T_m + N_{D_i} \times T^e_{D_i}$
  else
    if $W^e_{D_i} > 0$ **then**
      Calculate $S_{D_i} = W_{D_i}/T^e_{D_i}$
      Put S_{D_i} to *S Table*
    end if
  end if
end for
Sort *S Table* by S_{D_i} in a descending order
Return *S Table*, T_m

V. EXPERIMENT AND RESULT

We illustrated our experimental evaluations in this section. Section 6.1 represented our experimental configurations. Section 6.2 provided partial experimental results.

6.1 Experimental Configurations

We established the experimental environment based on our laboratorial setting. First, we developed a simulator in order to simulate multiple approaches and evaluate their performances. We selected Advanced Encryption Standard (AES) as the encryption method in our experiments. Next, cloud environments were set up in our lab as well. The hardware configuration was processor (Intel(R) Core(TM) i5-750 CPU @8M Cache, 2.66 GHz), 16 GB physical memory. The cloud computing side was installed on an Oracle VM Virtual Box workstation with running an Ubuntu 16.4.04 LTS Server.

Furthermore, we evaluated our DDES model's performances in two perspectives, which included both privacy protection level and computing efficiency. The experimental settings were in line with these two aspects, which consisted of a series of settings for reaching the expected evaluation goals. We simulated different workload sizes of the data transmission by configuring distinct amounts of data packages and the execution time of data encryption for each data package. In order to evaluate the privacy protection capability, we valued different data packages with different privacy weights. Moreover, we used a Brute Force (BF) algorithm to gain the optimal solution results in order to assess the performance of optimal solution acquisitions.

- **Setting 1:** We evaluated the data package amounts between 1 and 7. The encryption time for each data package wasbetween 5 and 10 units time. The privacy weight value for each data package was between 1 and 6. The no encryption execution time was configured between 1 and 5 units time.
- **Setting 2:** We evaluated the data package amounts between 1 and 5. The encryption time for each data package was between 2 and 15 units time. The privacy weight value for each data package was between 1 and 6. The no encryption execution time was configured between 1 and 5 units time.
- **Setting 3:** We evaluated the data package amounts between 3 and 11. The encryption time for each data package was between 1 and 50 units time. The privacy weight value for each data package was between 1 and 3. The no encryption execution time was configured between 1 and 3 units time.
- **Setting 4:** We evaluated the data package amounts between 3 and 6. The encryption time for each data package was between 10 and 15 units time. The privacy weight value for each data package was between 3 and 6. The no encryption execution time was configured between 1 and 2 units time.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

The purpose of evaluating the proposed method under different experimental settings was to examine DDES's performances while the input data were various and the constraints were varied. A few evaluations methods were used in our evaluations. First, we examined the performance differences between DDES and the optimal solutions under the configured settings. Two parameters were investigated, which were total weights and total execution time. Moreover, we compared the method generation time among DDES, dynamic programming and general BF while different experimental settings were operated. Finally, we analysed the privacy protection performance by comparing DDES with the optimal solutions.

Rp referred to the Performance Rate that displayed the performance of the target objective. V(DDES) denoted the variable value evaluated for DDES. V(Optimal) denoted the value of the optimal solution. Therefore, DDES performed an optimal solution when Rp = 1.

6.2 Experimental Results

We illustrated a few experimental results in this section. Fig. 3 and 4 represented a group of comparison results concerning the total PWV and the required execution time between our DDES and optimal solutions under Setting 1, respectively. Fig. 3 displayed that our approach had a similar performance to the optimal solutions in acquiring total privacy weight and Fig. 4 displayed the differences of the estimated execution time. Figures illustrated the results from the same experiment rounds. Most P values obtained from DDES were close to the optimal results that were obtained by BF algorithm.

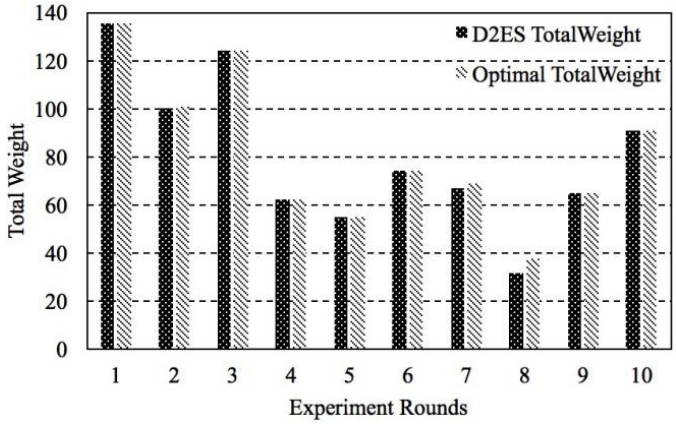

Fig. 3: Comparisons of total privacy weights between DDES and optimal solution under Setting 1.

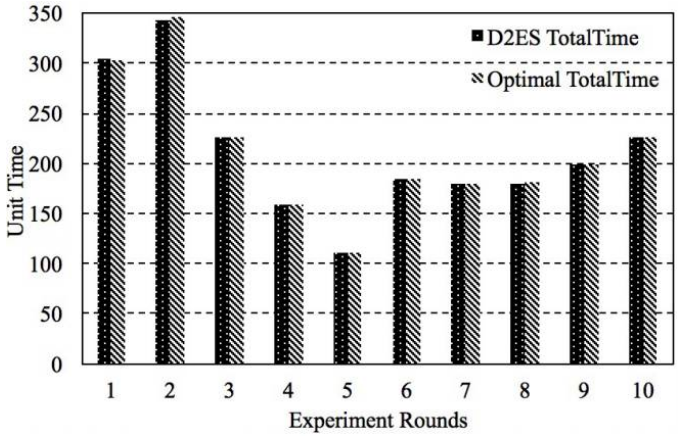

Fig. 4:Comparisons of total required execution time between DDES & optimal solution pairing with Fig. 3 under Setting 1.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 7, Special Issue 6, May 2018

2nd National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"

13th-14th March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

In summary, our DDES had better performances under different experimental settings. The experimental results reached our design goal and matched the theoretical results. Future research will address the practical measurements in real-world evaluations.

VI. CONCLUSION

This paper solely focused on the privacy issues of big data and considered the practical implementations in cloud computing. The proposed approach, DDES, was designed to maximize the efficiency of privacy protections. Main algorithm supporting DDES model was DED algorithm that was developed to dynamically alternative data packages for encryptions under different timing constraints. The experimental evaluations showed the proposed approach had an adaptive and superior performance.

VII. ACKNOWLEDGEMENT

We would to like to thank our HOD and project guide for helping us to find out the required resources in order to complete this project work and we would also like to thank all the professional bodies out there to who helped us consistently in our endeavours.

REFERENCES

[1] SCLPV: Secure Certificateless Public Verification for Cloud-Based Cyber-Physical-Social Systems Against Malicious Auditors Yuan Zhang, Student Member, IEEE, ChunxiangXu, Member, IEEE, Shui Yu, Senior Member, IEEE, Hongwei Li, Member, IEEE, and Xiaojun Zhang.

[2] Privacy-Preserving Public Auditing for Secure Cloud Storage Cong Wang, Student Member, IEEE, Sherman S.-M. Chow, Qian Wang, Student Member, IEEE, KuiRen, Member, IEEE, and Wenjing Lou, Member, IEEE.

[3] M. Maffei, G. Malavolta, M. Reinert, and D. Schroder. Privacy and access control for outsourced personal records. In IEEE Symposium on Security and Privacy, pages 341–358, San Jose, CA, USA, 2015. IEEE.

[4] C. Mulliner, W. Robertson, and E. Kirda. Hidden GEMs: Automated discovery of access control vulnerabilities in graphical user interfaces. In IEEE Symposium on Security and Privacy, pages 149–162, San Jose, CA, USA, 2014. IEEE.

[5] S. Sen, S. Guha, A. Datta, S. Rajamani, J. Tsai, and J. Wing. Bootstrapping privacy compliance in big data systems. In IEEE Symposium on Security and Privacy, pages 327–342, San Jose, CA, USA, 2014. IEEE.

[6] L. Zhu, Z. Hu, J. Heidemann, D. Wessels, A. Mankin, and N. Somaiya.Connection-oriented DNS to improve privacy and security. In IEEE Symposium on Security and Privacy, pages 171–186, San Jose, CA, USA, 2015. IEEE.