International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 9, September 2014)

Security Guidance for Critical Areas in Cloud and Solution of Dense Systems of Linear Equations

P.Madhuri
M.Tech Scholar -CSE
Sri Indu College of Engg and Tech
Ibrahimpatan, Hyderabad, TS, India

Gurram Bhavani
M.Tech Scholar-CSE Dept
Sri Indu College of Engg and Tech
Ibrahimpatan, Hyderabad, TS, India

C Shanker
Assistant Professor, Dept of CSE
Sri Indu College of Engg and Tech
Ibrahimpatan, Hyderabad, TS, India

Abstract: **Cloud computing economically enables customers with limited computational resources to outsource large-scale computations to the cloud. However, how to protect customers' confidential data involved in the computations then becomes a major security concern. In this paper, we present a secure outsourcing mechanism for solving large-scale systems of linear equations (LE) in cloud. Because applying traditional approaches like Gaussian elimination or LU decomposition (aka. direct method) to such large- scale LEs would be prohibitively expensive, we build the secure LE outsourcing mechanism via a completely different approach—iterative method, which is much easier to implement in practice and only demands relatively simpler matrix-vector operations. Specifically, our mechanism enables a customer to securely harness the cloud for iteratively finding successive approximations to the LE solution, while keeping both the sensitive input and output of the computation private. For robust cheating detection, we further explore the algebraic property of matrix-vector operations and propose an efficient result verification mechanism, which allows the customer to verify all answers received from previous iterative approximations in one batch with high probability. Thorough security analysis and prototype experiments on Amazon EC2 demonstrate the validity and practicality of our proposed design.**

***Index Terms*—Confidential data, computation outsourcing, system of linear equations, cloud computing**

1. LITERATURE SURVEY

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy n company strength. Once these things r satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system.

Cloud Computing: Cloud computing is the provision of dynamically scalable and often virtualized resources as a services over the internet Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them. Cloud computing represents a major change in how we store information and run applications. Instead of hosting apps and data on an individual desktop computer, everything is hosted in the "cloud"—an assemblage of computers and servers accessed via the Internet.

Cloud computing exhibits the following key characteristics:

- Agility improves with users' ability to re-provision technological infrastructure resources.
- Cost is claimed to be reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure. This is purported to lower barriers to entry, as infrastructure is typically provided by a third-party and does not need to be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is fine-grained with usage-based options and fewer IT skills are required for implementation. The e-FISCAL project's state of the art repository contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.
- Virtualization technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.
- Multi tenancy enables sharing of resources and costs across a large pool of users thus allowing for:
- Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
- Utilization and efficiency improvements for systems that are often only 10–20% utilized.
- Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.
- Performance is monitored and consistent and loosely coupled architectures are constructed using web services as the system interface.
- Security could improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. However, the complexity of security is greatly increased when data is distributed over a wider

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 9, September 2014)

area or greater number of devices and in multi-tenant systems that are being shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

- Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

2. INTRODUCTION

Cloud computing economically enables customers with limited computational resources to outsource large-scale computations to the cloud. However, how to protect customers' confidential data involved in the computations then becomes a major security concern. In this paper, we present a secure outsourcing mechanism for solving large-scale systems of linear equations (LE) in cloud. Because applying traditional approaches like Gaussian elimination or LU decomposition (aka. direct method) to such large- scale LEs would be prohibitively expensive, we build the secure LE outsourcing mechanism via a completely different approach—iterative method, which is much easier to implement in practice and only demands relatively simpler matrix-vector operations. Specifically, our mechanism enables a customer to securely harness the cloud for iteratively finding successive approximations to the LE solution, while keeping both the sensitive input and output of the computation private. For robust cheating detection, we further explore the algebraic property of matrix-vector operations and propose an efficient result verification mechanism, which allows the customer to verify all answers received from previous iterative approximations in one batch with high probability. Thorough security analysis and prototype experiments on Amazon EC2 demonstrate the validity and practicality of our proposed design.

We formulate the problem in the computation outsourcing model for securely solving large-scale systems of LE via iterative methods, and provide the secure mechanism design which fulfills input/output privacy, cheating resilience, and efficiency. Our mechanism brings computational savings as it only incurs O(n) local computation burden for the customer within each algorithm iteration and demands no unrealistic IO cost, while solving large scale LE locally usually demands more than O(n2) computation cost in terms of both time and memory requirements. We explore the algebraic property of matrix-vector multiplication to design a batch result verification mechanism, which allows customers to verify all answers computed by cloud from previous iterations in one batch, and further ensures both the efficiency advantage and the robustness of the design. Applying ordinary encryption techniques to the sensitive information before outsourcing could be one way to combat the security concern; it also makes the task of computation over encrypted data in general a very difficult problem. The cloud are not transparent enough to customers, no guarantee is provided on the quality of the computed results from the cloud possible software/hardware malfunctions and/or outsider attacks might also affect the quality of the computed results. Thus, we argue that the cloud is intrinsically not secure from the viewpoint of customers.

In existing approaches and the computational practicality motivates us to design secure mechanism of outsourcing LE via a completely different approach iterative method, where the solution is extracted via finding successive approximations to the solution until the required accuracy is obtained. Compared to direct method, iterative method only demands relatively simpler matrix-vector operations, which is much easier to implement in practice and widely adopted for large-scale LE. To the best of our knowledge, no existing work has ever successfully tackled secure protocols for iterative methods on solving large-scale systems of LE in the computation outsourcing model.

Applying ordinary encryption techniques to the sensitive information before outsourcing could be one way to combat the security concern; it also makes the task of computation over encrypted data in general a very difficult problem. The cloud are not transparent enough to customers, no guarantee is provided on the quality of the computed results from the cloud possible software/hardware malfunctions and/or outsider attacks might also affect the quality of the computed results. Thus, we argue that the cloud is intrinsically not secure from the viewpoint of customers. The execution time of a computer program depends not only on the number of operations it must execute, but on the location of the data in the memory hierarchy, solving such large-scale problems on customer's weak computing devices can be practically impossible, due to the inevitably involved huge IO cost.

3. PROPOSED SYSTEM

We propose a very efficient cheating detection mechanism to effectively verify in one batch of all the computation results by the cloud server from previous algorithm iterations with high probability. We formulate the problem in the computation outsourcing model for securely solving large-scale systems of LE via iterative methods, and provide the secure mechanism design which fulfills input/output privacy, cheating resilience, and efficiency. We explore the algebraic property of matrix-vector operations to design a batch verification mechanism, which allows customers to verify all results of previous iterations from cloud in one batch. It ensures both the efficiency advantage and robustness of the design.

Fully homomorphic encryption (FHE) scheme, a general result of secure computation outsourcing has been shown viable in theory, where the computation is represented by an encrypted combinational Boolean circuit that allows to be evaluated with encrypted private inputs.

Advantages

- The problem of securely outsourcing large-scale systems of LE via iterative methods, and provide mechanism

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 9, September 2014)

designs fulfilling input/output privacy, cheating resilience, and efficiency.
- Our mechanism brings computational savings.
- We explore the algebraic property of matrix-vector operations to design a batch verification mechanism, which allows customers to verify all results of previous iterations from cloud in one batch. It ensures both the efficiency advantage and robustness of the design.

Limitations:

- Applying ordinary encryption techniques to the sensitive information before outsourcing could be one way to combat the security concern; it also makes the task of computation over encrypted data in general a very difficult problem
- The cloud are not transparent enough to customers, no guarantee is provided on the quality of the computed results from the cloud possible software/hardware malfunctions and/or outsider attacks might also affect the quality of the computed results. Thus, we argue that the cloud is intrinsically not secure from the viewpoint of customers.
- The execution time of a computer program depends not only on the number of operations it must execute, but on the location of the data in the memory hierarchy, solving such large-scale problems on customer's weak computing devices can be practically impossible, due to the inevitably involved huge IO cost.

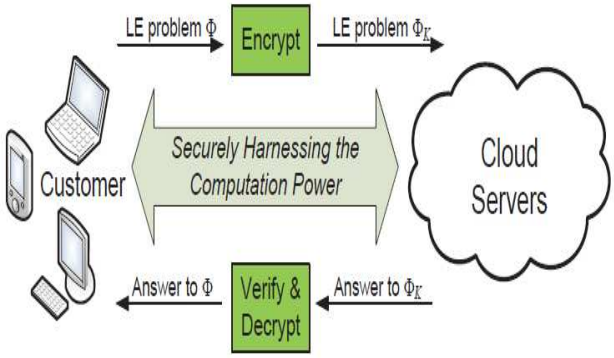

Fig 1: Content Diagram of Project

4. DESIGN

Design is a meaningful engineering representation of something that is to be built. Software design is a process through which the requirements are translated into a representation of the software. Design is the place where quality is fostered in software engineering. Design is the perfect way to accurately translate a customer's requirement in to a finished software product. Design creates a representation or model, provides detail about software data structure, architecture, interfaces and components that are necessary to implement a system. This chapter discusses about the design part of the project. Here in this document the various UML diagrams that are used for the implementation of the project are discussed.

Design Principle: The Unified Modeling Language (UML) is a visual modeling language used to specify, visualize, construct and document a software intensive system. The embedded real-time software systems encountered in applications such as telecommunications, school systems, aerospace, and defense typically tends to be large and extremely complex. It is crucial in such systems that the software is designed with a sound architecture. A good architecture not only simplifies construction of the initial system, but also, readily accommodates changes forced by a steady stream of new requirements. The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects. Using the UML helps project teams communicate, explore potential designs, and validate the architectural design of the software. The primary goals in the design of the UML are: Provide users with a ready-to-use, expressive visual modeling language so they can develop and exchange meaningful models. Provide extensibility and specialization mechanisms to extend the core concepts. Be independent of particular programming languages and development processes. Provide a formal basis for understanding the modeling language. Encourage the growth of the OO tools market. Support higher-level development concepts such as collaborations, frameworks, patterns and components. Integrate best practices.

5. MODULE DESIGN

- Cloud Computing
- Homomorphic Encryption
- General Techniques

Cloud Computing: Cloud computing is the provision of dynamically scalable and often virtualized resources as a services over the internet Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them. Cloud computing represents a major change in how we store information and run applications. Instead of hosting apps and data on an individual desktop computer, everything is hosted in the "cloud"—an assemblage of computers and servers accessed via the Internet.
Homomorphic Encryption: An efficient semantically-secure encryption scheme with additive homomorphic property.Given two integers x1 and x2, we have Enc(x1 + x2) = Enc(x1)_Enc(x2), and also Enc(x1_x2) = Enc(x1)x2 . In our implementation we adopt the one presented by Paillierin . The Paillier cryptosystem is a publickey cryptosystem.

General Techniques

ProbTransform: In this phase, cloud customer would initialize a randomized key generation algorithm and prepare

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 9, September 2014)

the LE problem into some encrypted form _K via key K for phase ProbSolve. Transformation and/or encryption operations will be needed when necessary.

ProbSolve: In this phase, cloud customer would use the encrypted form _K of LE to start the computation outsourcing process. In case of using iterative methods, the protocol ends when the solution within the required accuracy is found.

Result Verify: In this phase, the cloud customer would verify the encrypted result produced from cloud server, using the randomized secret key K. A correct output x to the problem is produced by decrypting the encrypted output. When the validation fails, the customer outputs, indicating the cloud server was cheating.

6. IMPLEMENTATION

The most crucial phase of any project is the implementation. This includes all those activities that take place to convert from the old system to the new system. It involves setting up of the system for use by the concerned end user. A successful implementation involves a high level of interaction between the analyst, programmers and the end user. The most common method of implementation is the phased approach, which involves installation of the system concurrently with the existing system. This has its advantage in that the normal activity carried out, as part of the existing system is anyway hampered. The end users are provided with sufficient documentation and adequate training in the form of demonstration/presentation in order to familiarize with the system.

Process Model Used With Justification

SDLC (Umbrella Model):

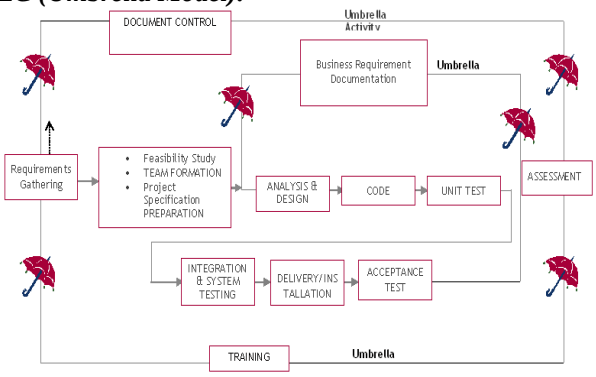

Fig 2. Umbrella Model

SDLC is nothing but Software Development Life Cycle. It is a standard which is used by software industry to develop good software.

Stages in SDLC:

- Requirement Gathering
- Analysis
- Designing
- Coding
- Testing
- Maintenance

Requirements Gathering stage: The requirements gathering process takes as its input the goals identified in the high-level requirements section of the project plan. Each goal will be refined into a set of one or more requirements. These requirements define the major functions of the intended application, define operational data areas and reference data areas, and define the initial data entities. Major functions include critical processes to be managed, as well as mission critical inputs, outputs and reports. A user class hierarchy is developed and associated with these major functions, data areas, and data entities. Each of these definitions is termed a Requirement. Requirements are identified by unique requirement identifiers and, at minimum, contain a requirement title and textual description.

7. OUTPUT SCREENS

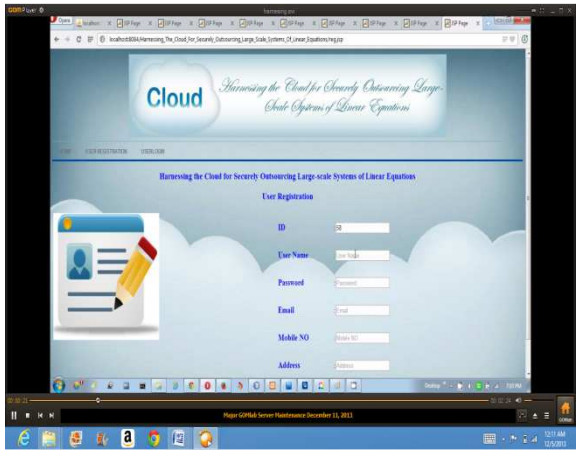

Fig.3. SS 2: User Registration Form

Fig.4. SS 3: User Login Page

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 9, September 2014)

Fig.5. SS 4: File uploading page

Fig.6. SS 5: Get the key while download a file

8. CONCLUSION

In this paper, we investigated the problem of securely outsourcing large-scale LE in cloud computing. Different from previous study, the computation outsourcing framework is based on iterative methods, which has the benefits of easy-to-implement and less memory requirement in practice. This is especially suitable for the application scenario, where computational constrained customers want to securely harness the cloud for solving large-scale problems. We also investigated the algebraic property of the matrix-vector multiplication and developed an efficient and effective cheating detection scheme for robust result verification. Thorough security analysis and extensive experiments on the real cloud platform demonstrate the validity and practicality of the proposed mechanism.

REFERENCES

[1]. A. Arasu and H. Garcia-Molina, "Extracting Structured Data from Web Pages," Proc. SIGMOD Int'l Conf. Management of Data, 2003.

[2]. L. Arlotta, V. Crescenzi, G. Mecca, and P. Merialdo, "Automatic Annotation of Data Extracted from Large Web Sites," Proc. Sixth Int'l Workshop the Web and Databases (WebDB), 2003.

[3]. P. Chan and S. Stolfo, "Experiments on Multistrategy Learning by Meta-Learning," Proc. Second Int'l Conf. Information and Knowledge Management (CIKM), 1993.

[4]. W. Bruce Croft, "Combining Approaches for Information Retrieval," Advances in Information Retrieval: Recent Research from the Center for Intelligent Information Retrieval, Kluwer Academic, 2000.

[5]. V. Crescenzi, G. Mecca, and P. Merialdo, "RoadRUNNER: Towards Automatic Data Extraction from Large Web Sites," Proc. Very Large Data Bases (VLDB) Conf., 2001.

[6]. S. Dill et al., "SemTag and Seeker: Bootstrapping the Semantic Web via Automated Semantic Annotation," Proc. 12th Int'l Conf./ World Wide Web (WWW) Conf., 2003.

[7]. H. Elmeleegy, J. Madhavan, and A. Halevy, "Harvesting Relational Tables from Lists on the Web," Proc. Very Large Databases (VLDB) Conf., 2009.

[8]. D. Embley, D. Campbell, Y. Jiang, S. Liddle, D. Lonsdale, Y. Ng, and R. Smith, "Conceptual-Model-Based Data Extraction from Multiple-Record Web Pages," Data and Knowledge Eng., vol. 31,no. 3, pp. 227-251, 1999.

[9]. D. Freitag, "Multistrategy Learning for Information Extraction," Proc. 15th Int'l Conf. Machine Learning (ICML), 1998.

About the authors:

P.Madhuri: Completed B.Tech from Jyothishmathi Institute of Technological Sciences, karimnagar, JNTUH with Distinction.Currently pursuing M.Tech 2nd Year in Sri Indu College of Engineering and Technology. Areas of Interest are Web Technologies and Data Base Management Systems, Object Oriented Analysis& Design.

Gurram Bhavani: Completed B.Tech from D.V.R College of Engineering and Technology, Hyderabad, JNTUH with Distinction.Currently pursuing M.Tech 2nd Year in Sri Indu College of Engineering and Technology. Areas of Interest are Cloud Computing, Web Technologies and Data Base Management Systems, Object Oriented Analysis& Design.

C SHANKER, Currently working as a Assistant Professor in Sri Indu College of Engg & Technology. Having 6 years' experience. Studied B.Tech in Vijaya rural Engg College, Nizamabad. M.Tech in J.B.Institute of Engineering & Technology, Hyderabad. Area of Interest are Cloud Computing, Data Mining, Software Engineering, Network Security.