UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES AND ACADEMIC ADVANCEMENT

Query Privacy Preservability in Cloud Computing

A Thesis

Submitted to the College of Graduate Studies and Academic Advancements in Partial Fulfillment of the Requirement for the Degree of Master of Science in Information Technology

Prepared By:

Suliman Ali Dambas Adam

Supervisor By:

Prof. Noureldien Abdelrhman Noureldien

Jan 2017

Abstract

Cloud Computing has been defined as a new business model. It is an emerging paradigm because of that it has been given high attention from many researchers. The main advantage of cloud computing is to reduce the cost of computing while, at the same time, the main disadvantage of using cloud computing is the lack of query privacy information. The query privacy in the cloud computing is more critical. To protect this privacy done by index privacy, keyword privacy, and trapdoor unlinkability, using some schemes and algorithms to secure user in the cloud, such as range queries and keyword queries.

Range query schemes by use bucketing schemes to preserve the query privacy over encrypted data, But it leaks some information for user query, this problem is solved by some proposed algorithms. Order Preserving Encryption (OPE) is an encryption algorithm that used to preserve the natural ordering of the plaintexts. OPE scheme solves partially the problem of searching over encrypted data.

Keyword query schemes is an encryption algorithms that used to preserve the query keyword privacy over encrypted data by using Asymmetric searchable encryption, Symmetric searchable encryption and Adaptively secure searchable symmetric encryption.

This thesis was used schemes and algorithms to preserve the query privacy of user data. It has been investigated encryption algorithms that preserve privacy of query such as keyword query.

This research surveys all schemes that have been proposed to preserved query privacy in cloud environment.

المستخلص

تعتبرالحوسبة السحابية من اهم المواضيع التي حصلت علي اهتمام كبيرا من قبل عديد من الباحثين. تتميز الحوسبة السحابية بميزة اساسية وهي انخفاض تكاليف الحوسبة و في الوقت نفسة احدي العيوب الحوسب السحابية هو تسريب خصوصية معلومات الاستعلام الذي يعتبر من اهم محاور في الحوسبة السحابية. لحماية هذه الخصوصية الذي تتم بواسطة خصوصية مؤشر والخصوصية الكلمات الرئيسية، والقدرة الباب المسحور لإلغاء الربط، وذلك باستخدام بعض برامج وخوارزميات لتأمين المستخدم في السحابة، مثل الاستفسارات نطاق والاستفسارات الكلمات الرئيسية.

مدى نظم الاستعلام التي تستخدم للحفاظ علي خصوصية الاستعلام في حوسبة السحابة باستخدام نظام تقسيم البيانات في الاستعلام الي فترات غير متداخلة ولكنه تسرب بعض المعلومات عن الاستعلام في حالة تداخل الفترات ، هذه المشكلة تم حلها عن طريق بعض الخوارزميات المقترحة.

(OPE) هو خوارزمية التشفير التي تستخدم للحفاظ على طبيعة ترتيب النصوص العادية في النصوص المشفرة. (OPE) يحل بعض مشاكل البحث علي البيانات المشفرة.

باستخدام بعض الخوارميات التشفير مثل المتماثل و غير المتماثل علي الكلمات الرئيسية المعنيه في الاستعلام هي .خوارزميات التشفير التي تستخدم للحفاظ على خصوصية الكلمات الاستعلام على البيانات المشفرة

وقد تم .وقد استخدمت هذه الأطروحة المخططات والخوارزميات للحفاظ على خصوصية الاستعلام بيانات المستخدم التحقيق فيها خوارزميات التشفير التي تحافظ على خصوصية الاستعلام مثل الاستعلام الكلمات المعينه.

يستعرض هذا البحث كل المخططات التي تم اقتراحها في الخصوصية الاستعلام في الحفاظ على البيئة السحابية.

1.1 Introduction

The concept of cloud computing has recently emerged as a model for hosting and delivering services via the Internet. Cloud computing has made a big shift in the field of information technology, where this model allows the user to access services and files from anywhere and at any time. That service provided by provider called cloud services provider that rents out these services to customers (individuals or companies). This rent depending on the type of service required by the client.

Privacy refers to sensitive information about one person or a group that is expected to be secluded or hidden from others (e.g., identity, address, health, and hobbies). When users access cloud data and use cloud data services, it is necessary to preserve their privacy. Many users pay more attention to their privacy protection when they access cloud data or use cloud services. In particular, they expect to hide their identity while using cloud data services. Some users also want their operations on the data and the information retrieved from a cloud to be properly protected. For instance, the keywords queried over the outsourced data and the query results returned by a cloud should not be exposed to others. Moreover, it is expected that users' access behaviors and habits should not be inferred by any other parties in cloud.

User privacy in cloud services includes identity privacy, query privacy, and access pattern privacy. Different privacy issues can be addressed with different protection techniques and need some specific considerations. In this thesis we will address only query privacy. Query privacy can be effectively implemented by using secure indexes, virtual/dummy keywords, or random trapdoors [1].

1.2 Problem Statement

Cloud computing is one of the recent technologies and provides many services to users. Search over encrypted data is a challenging problem in the cloud privacy field. Protect query privacy is a fundamental issue in the cloud computing paradigm. The traditional encryption techniques are preventing unauthorized access to sensitive data while at the same time do not preserve the query privacy. Query privacy schemes solve the problem of searching over encrypted data partially, but it leaks some information.

1.3 Research Questions

Goals of this thesis have been accomplished by answering the following research questions:

1- What is the query privacy in cloud issue?

2- What are the solutions that have been proposed to achieve protection of query privacy?

3- How can we classify these proposed schemes?

1.4 Research Objectives

This research aims to study query privacy preservation over encrypted cloud data, and survey solutions of how to protect query privacy in cloud, to provide readers with an organized up to date literature on the topic.

1.5 Motivations

Because moving data to cloud environment means that control of the data is becoming under third party at remote server, customers need to ensure that their queries was maintained safely and no any compromise was happen to it. To ensure protection to query privacy we will focus in technique and solutions proposed by researchers for this issue.

1.6 Research Methodology

In this research our methodology is to look at query privacy issue in cloud data. After that we go to classified scientific papers that proposed solutions to these privacy problems. And our methodology in collecting scientific papers is depending on two major factors. First factor is source of papers, we collecting papers from best two sources in Information Technology field (ACM and IEEE) and popular source (Google Scholar) and Journal related with query privacy in cloud field. Second factor is date of published papers, we focusing at modern papers that published in last six years from 2016.

A survey of literature methodology is used to answer the research questions. Chapter (3) gives details of the research methodology.

1.7 Research Scope

This research thesis is concerned only to study methods dedicated to protect query privacy on encrypted cloud data.

1.8 Thesis contribution

Thesis contributions are:

1. Identifying the query privacy issue with proposed protection solutions.
2. Classifying schemes that have been proposed to the protect query privacy.

1.9 Thesis Structure

The remaining contents of the thesis are organized as follows:

Chapter 2 contains important aspects of cloud computing, including its definition, essential characteristics, models of cloud services, pros and cons of using cloud computing. The chapter also briefly stated the challenges and security issues that faces cloud computing. Chapter 3 presents the research methodologies being used to conduct this research. Chapter 4 discusses query privacy issues and surveys the proposed solutions to preserve privacy. Chapter 5 states the conclusions and recommendations for future works.