IMPROVED SECURE ANTI-COLLUSION DATA SHARING SCHEME FOR DYNAMIC GROUPS IN THE CLOUD

K. Gunasekaran


*Corresponding Author: S.Raghul,
Phone: +91-4294-250242, 250220 ; Fax: +91-4294-250219
E-mail: kgunamgp@gmail.com, cseraghul@gmail.com

Received: 01/01/2017, Revised: 17/03/2017 and Accepted: 01/04/2017

Abstract

Cloud data sharing and low maintenance, provides a better computing, with the characteristics of intrinsic utilization of resources. In cloud computing, cloud service providers offer an abstraction of infinite storage space for clients to host data. It can help clients reduce their financial overhead of data managements by migrating the local managements system into cloud servers. Propose a secure data sharing scheme, which can achieve secure key distribution and data sharing for dynamic group. In this project, existing scheme is able to support dynamic groups efficiently, when a new user joins in the group or a user is revoked from the group, the private keys of the other users do not need to be recomputed and updated. Moreover, scheme can achieve secure user revocation; the revoked users can not be able to get the original data files once they are revoked even if they conspire with the untrusted cloud. To avoid these disadvantages, Data Anti collusion is a technique for eliminating duplicate copies of data, and has been widely used in cloud storage to reduce storage space and upload bandwidth and secure cloud data stored. The proposed convergent encryption model has been extensively adopted for secure Anti collusion and to efficiently and reliably manage a huge number of convergent keys.

Keywords: Migrating, revocation, Covergent Keys, Elimination,

1. Introduction

Cloud computing is internet-based computing in which large groups of remote servers are networked to allow sharing of data-processing tasks, centralized data storage, and online access to computer services or resources. Clouds can be classified as public, private or hybrid. Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. The main enabling technology for cloud computing is virtualization. Virtualization software allows a physical computing device to be electronically separated into one or more "virtual" devices, each of which can be easily used and managed to perform computing tasks. Cloud computing adopts concepts from Service oriented Architecture (SOA) that can help the user break these problems into services that can be integrated to provide a solution.
2. Cloud Computing Types

2.1. Private Cloud

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party, and hosted either internally or externally. Undertaking a private cloud project requires a significant level and degree of engagement to virtualize the business environment, and requires the organization to reevaluate decisions about existing resources. Self-run data centers are generally capital intensive. They have a significant physical footprint, requiring allocations of space, hardware and environmental controls.

2.2. Public Cloud

A cloud is called a "public cloud" when the services are rendered over a network that is open for public use. Public cloud services may be free or offered on a pay-per-usage model. However, security consideration may be substantially different for services (applications, storage, and other resources) that are made available by a service provider for a public audience and when communication is effected over a non-trusted network. Generally, public cloud
service providers like Amazon AWS, Microsoft and Google own and operate the infrastructure at their data center and access is generally via the Internet. AWS and Microsoft also offer direct connect services called "AWS Direct Connect" and "Azure Express Route" respectively, such connections require customers to purchase or lease a private connection to a peering point offered by the cloud provider.

2.3. Hybrid Cloud

Hybrid Cloud is an integrated cloud service utilizing both private and public clouds to perform distinct functions within the same organization. All cloud computing services should offer certain efficiencies to differing degrees but public cloud services are likely to be more cost efficient and scalable than private clouds.

- Separate cloud providers team up to provide both private and public services as an integrated service
- Individual cloud providers offer a complete hybrid package
- Organizations who manage their private clouds themselves sign up to a public cloud service which they then integrate into their infrastructure

Features of Hybrid Cloud

- Scalability
- Cost efficiencies
- Security
- Flexibility

2.4 Cloud Services

Although a cloud is a remotely accessible environment, not all IT resources residing within a cloud can be made available for remote access. For example, a database or a physical server deployed within a cloud may only be accessible by other IT resources that are within the same cloud. A software program with a published API may be deployed specifically to enable access by remote clients.

Network Cloud Services

What you may not have thought about is that every one of these consumer application cloud services uses network cloud services. In fact, the word “cloud” comes from the fact that many years ago those of us who built and sold client server applications, software and hardware used to draw a picture with the PC connected to a network and the network connected to a server. Since none of us actually understood how the network worked, we drew a cloud and labeled it “network” and left it at that. In those days companies built their own networks, but today consumers and businesses use network cloud services delivered by companies like AT&T, Verizon, Masergy and Sprint.

Application Cloud Services

So far we’ve focused on consumer application cloud services, but for the past ten years the fastest-growing business applications have all been delivered as cloud services. Since 1999, fifteen companies that deliver business application cloud services have become public companies. Some of these companies have been acquisition targets. An informal analysis of forty of the Fortune 100 showed only two companies that did not have at least one of these applications running. Of course today, nearly all traditional application software companies like Oracle and JDA offer to manage their applications as a service.
Platform Cloud Services

This brings us to the last group of cloud services. Platform cloud services are used by software developers to build new applications and by operations managers to manage their application, compute and storage cloud services. Horizontal platforms offer a great deal of flexibility, but if you know you want to leverage NetSuite’s schema to build an MRP application, like Rootstock did, then choosing a particular vertical app can significantly speed the development and reduce the cost to build new applications. Platform cloud services also provide the operations management specialists with a range of services.

Bandwidth Requirements

If you are going to adopt the cloud framework, bandwidth and the potential bandwidth bottleneck must be evaluated in your strategy. Virtualization implementers found that the key bottleneck to virtual machine density is memory capacity; now there's a whole new slew of servers coming out with much larger memory footprints, removing memory as a system bottleneck. Cloud computing negates that bottleneck by removing the issue of machine density from the equation sorting that out becomes the responsibility of the cloud provider, freeing the cloud user from worrying about it. For cloud computing, bandwidth to and from the cloud provider is a bottleneck.

A blade server is a server that has been optimized to minimize the use of physical space and energy. One of the huge advantages of the blade server for cloud computing use is bandwidth speed improvement. For example, the IBM BladeCenter is designed to accelerate the high-performance computing workloads both quickly and efficiently. Just as the memory issue had to be overcome to effectively alleviate the bottleneck of virtual high machine density, the bottleneck of cloud computing bandwidth must also be overcome, so look to the capabilities of your provider to determine if the bandwidth bottleneck will be a major performance issue.

Advantages

The pros of cloud computing are obvious and compelling. Do you really want them cluttering your expensive computers with their personal emails, illegally shared MP3 files, and naughty YouTube videos when you could leave that responsibility to someone else? Cloud computing allows you to buy in only the services you want, when you want them, cutting the upfront capital costs of computers and peripherals. You avoid equipment going out of date and other familiar IT problems like ensuring system security and reliability. You can add extra services at a moment's notice as your business needs change. It's really quick and easy to add new applications or services to your business without waiting weeks or months for the new computer (and its software) to arrive.

3.0 System Analysis
3.1 Problem Definition

The existing system maintained by the cloud service providers, provides storage space for hosting data files in a pay-as-you-go manner. However, the cloud is un-trusted since the cloud service providers are easily to become un-trusted. Therefore, the cloud will try to learn the content of the stored data. Group manager takes charge of system parameters generation, user registration, and user revocation.
In the existing applications, the group manager usually is the leader of the group. Therefore, the research assumes that the group manager is fully trusted by the other parties. Group members (users) are a set of registered users that will store their own data into the cloud and share them with others. In the scheme, the group membership is dynamically changed, due to the new user registration and user revocation.

However, this existing scheme is not secure because of the weak protection of commitment in the phase of identity token issuance. The private keys of the other users do need to be recomputed and updated. In general, scheme cannot achieve secure key distribution, fine access control and secure user revocation in the dynamic cloud group environment.

3.2 Existing System

In this project existing system maintained by the cloud service providers, provides storage space for hosting data files in a pay-as-you-go manner. However, the cloud is untrusted since the cloud service providers are easily to become untrusted. Therefore, the cloud will try to learn the content of the stored data. Group manager takes charge of system parameters generation, user registration, and user revocation.

In the existing applications, the group manager usually is the leader of the group. Therefore, we assume that the group manager is fully trusted by the other parties. Group members (users) are a set of registered users that will store their own data into the cloud and share them with others. In this scheme, the group membership is dynamically changed, due to the new user registration and user revocation. The existing system includes an attribute-based access control scheme using CP-ABE with efficient attribute and user revocation capability for data outsourcing systems. The proposed scheme has following advantages with regard to the security and scalability compared to the previous revocable CP-ABE schemes.

In existing system, first, enabling user access control enhances the backward/forward secrecy of outsourced data on any membership changes in attribute groups compared to the attribute revocation schemes. Second, the user access control can be done on each attribute level rather than on system level, so that more fine-grained user access control can be possible.

Drawbacks of Existing System

- The data owner need to take full charge of maintaining all the membership lists for each attribute group to enable the direct user revocation.
- Keys are assigned randomly and independently from each other.
- All the data is maintained by single service provider.
- The single data service manager is in charge of managing the attribute group keys per each attribute group.
- All the nodes are treated equally and weak capable nodes also require huge computations.
- All the mirror nodes store the file with same encryption mechanism.
- Unauthorized data leakage still remains a problem due to the potential exposure of decryption keys.
- Only single cloud provider environment is considered.

3.3 Proposed System

The proposed system implements all the existing system concepts in which the Cipher text-Policy Attribute-
Based Encryption with User Revocation is carried out. Like existing system, the proposed scheme also adapts a dual encryption approach to overcome the user access control problem in attribute-based encryption system. In addition, multiple service providers are included and data is distributed among them. User privileges may be varying for data maintained by different service providers. This requires different kind of encryption mechanisms in data maintained by different service providers and so computation overhead is reduced.

**Advantages**

- Any service provider may revoke users if unauthorized user tries to access the data above a given count.
- Data servicing is maintained by more than one service provider.
- All data service manager take charge of managing the attribute group keys per each attribute group.
- Keys are assigned based on a condition and unique among all users.
- Partial data of files are taken from multiple mirror locations and send to selected client.
- Suitable for very large size files.
- Irrelevant size blocks of data are handled among the multiple cloud service providers based on their computational capabilities.
- Different trust level is set to different cloud providers and encryption or decryption is varied based on the clouds computational capability.

**4.0 System Methodology**

The data outsourcing scenario challenges the approaches of traditional access control architectures such as reference monitor, where a trusted server is in charge of defining and enforcing access control policies. This assumption no longer holds in modern data outsourcing systems, because users want to be able to share private contents with a group of people they selected and to define some access policy and enforce it on the contents. Thus, it is desirable to put the access policy decisions in the hands of the data owners.

The evaluation demonstrates that Dekey incurs limited overhead in normal upload/download operations in realistic cloud environments. This thesis study makes new construction Dekey to provide efficient and reliable convergent key management through convergent key de duplication and secret sharing. Dekey supports both file-level and block-level de duplications. Security analysis is demonstrates that Dekey is secure in terms of the definitions specified in the proposed security model. Symmetric encryption uses a common secret key to encrypt and decrypt information. Since the key used for this experimental work are very weak, the existing system is less secure. User revocation management is not implemented. The key can be management only within the group members.

**Setup**

Attribute setup algorithm is executed a randomized algorithm that takes no input other than the implicit security parameter. It outputs the public key PK and a master key MK.

**4.1 Attribute Key Generation**

Attribute key generation algorithm is executed which takes input the master key MK, a set of attributes \( \Lambda \subseteq L \), and a set of user indices \( Y \subseteq u \) as parameters. It outputs a set of private attribute keys SK for each user in U that
identifies with the attributes set.

4.2 Key Encrypting Key Generation

The key encrypting key (KEK) generation algorithm is executed in this module, which takes a set of user indices $U \subseteq \mathbb{u}$ as input, and outputs KEKs for each user in $U$, which will be used to encrypt attribute group keys $K_{A_{i}}$ for each $G_{i} \subseteq G$.

4.3 Encrypt

Attribute encryption algorithm (which is a randomized algorithm) that takes as input the public parameter $PK$, a message $M$, and an access structure ‘$A$’ over the universe of attributes. It outputs a cipher text $CT$ such that only a user who possesses a set of attributes that satisfies the access structure will be able to decrypt the message.

4.4 Re-Encrypt

Attribute re-encryption algorithm is a randomized algorithm that takes as input the ciphertext $CT$ including an access structure ‘$A$’, and a set of attribute groups $G$. If the attribute groups appear in ‘$A$’, it re-encrypts $CT$ for the attributes; else, returns. Specifically, it outputs a re-encrypted ciphertext $CT^{*}$ such that only a user who possesses a set of attributes that satisfies the access structure and has a valid membership for each of them at the same time will be able to decrypt the message.

4.5 Decrypt

Attribute decryption algorithm is executed which takes as input the ciphertext $CT^{*}$ which contains an access structure ‘$A$’, a private key $SK$, and a set of attribute group keys $K_{A_{i}}$ for a set of attributes $A$. The decryption can be done if $f_{A} \square A_{i}$ satisfies ‘$A$’ and $K_{A_{i}}$ is not revoked for any user.

4.6 Encryption and Decryption for Different service providers

The cloud services provider is data contains most important information and in order to protect the data security, more privileged service providers view most of the data and less privileged service providers view limited data.

4.7 Revocation of Users

In this module, consider the revocation of users in the given group. If the original (first) user of the group intimates the server with a user’s (B) revocation, then the server rejects the proof of ownership submitted by that user (B).

4.8 Add User

In this module, user id, username, mail id, password and random globally unique identifier is generated which will be used as the tag for further modules is added to ‘Users’ table.
SYSTEM DIAGRAM
PROPOSED SYSTEM

Encrypted Data (Data Owner)

SERVICES PROVIDER
- Data Server
- Data Services Manager

USER

Access

Manage

Trust Authority (TAP and BAP)

Manage

Encrypt ed Data
5.0 SYSTEM DESIGN

5.1 Input Design

Input design is the process of converting user-originated inputs to a computer understandable format. Input design is one of the most expensive phases of the operation of computerized system and is often the major problem of a system. A large number of problems with a system can usually be tracked backs to fault input design and method. Every moment of input design should be analyzed and designed with utmost care. The design of the input should be made the input as the over to the numerous networks in the reliable area that should be passed as the installation in the remote network. It has the following constraints in the input database.

- All the files from the disk should be acquired by data.
- It is suitable to more available data clearance and made available.
- The menu of design should be understandable and it is in the right format.
Add User

In this module, user id, username, mail id, password and random globally unique identifier is generated which will be used as the tag for further modules is added to ‘Users’ table.

5.2. Output Design

Output design generally refers to the results and information that are generated by the system for many end-users; output is the main reason for developing the system and the basis on which they evaluate the usefulness of the application the output is designed in such a way that it is attractive, convenient and informative. Forms are designed in C# .NET with various features, which make the console output more pleasing. As the outputs are the most important sources of information to the users, better design should improve the system’s relationships with user and also will help in decision-making. Form design elaborates the way of output is presented and the layout available for capturing information. The following forms are used for the output form.

- Key Encrypting Key Generation

5.3  Key Encrypting Key Generation

The key encrypting key (KEK) generation algorithm is executed in this module, which takes a set of user indices \( U \subseteq u \) as input, and outputs KEKs for each user in \( U \), which will be used to encrypt attribute group keys \( K_A \), for each \( G_i \subseteq G \).

5.4. SYSTEM REQUIREMENT

5.4.1. HARDWARE SPECIFICATION

This section gives the details and specification of the hardware on which the system is expected to work.

- Processor: Pentium IV 1.7 GHz
- Hard Disk Capacity: 40 GB
- RAM: 1 GB SD(DDR)
- Monitor: 15" Color
- Keyboard: 102 keys
- Mouse: Optical Mouse

5.4.2. SOFTWARE SPECIFICATION

This section gives the details of the software that are used for the development.

- Front-End: Visual C#. Net
- Web Technology: ASP.NET 2005
- Operating System: Windows XP SP2

5.3.3. FRONT END: THE .NET FRAMEWORK

The .NET Framework is a new computing platform that simplifies application development in the highly distributed environment of the Internet.
5.3.4. BACK END: FEATURES OF SQL-SERVER

Microsoft SQL Server is a relational model database server produced by Microsoft. Its primary query languages are T-SQL and ANSI SQL. The OLAP Services feature available in SQL Server version 7.0 is now called SQL Server 2000 Analysis Services. The term OLAP Services has been replaced with the term Analysis Services. Analysis Services also includes a new data mining component. The Repository component available in SQL Server version 7.0 is now called Microsoft SQL Server 2000 Meta Data Services. References to the component now use the term Meta Data Services. The term repository is used only in reference to the repository engine within Meta Data Services. Microsoft® SQL Server™ 2000 features include:

Internet Integration

The SQL Server 2000 database engine includes integrated XML support. It also has the scalability, availability, and security features required to operate as the data storage component of the largest Web sites. The SQL Server 2000 programming model is integrated with the Windows DNA architecture for developing Web applications, and SQL Server 2000 supports features such as English Query and the Microsoft Search Service to incorporate user-friendly queries and powerful search capabilities in Web applications.

Scalability and Availability

The same database engine can be used across platforms ranging from laptop computers running Microsoft Windows® 98 through large, multiprocessor servers running Microsoft Windows 2000 Data Center Edition. SQL Server 2000 Enterprise Edition supports features such as federated servers, indexed views, and large memory support that allow it to scale to the performance levels required by the largest Web sites.

Enterprise-Level Database Features

The SQL Server 2000 relational database engine supports the features required to support demanding data processing environments. The database engine protects data integrity while minimizing the overhead of managing thousands of users concurrently modifying the database.

SQL Server 2000 distributed queries allow you to reference data from multiple sources as if it were a part of a SQL Server 2000 database, while at the same time, the distributed transaction support protects the integrity of any updates of the distributed data. Replication allows to also maintaining multiple copies of data, while ensuring that the separate copies remain synchronized. The replicate a set of data to multiple, mobile, disconnected users, have them work autonomously, and then merge their modifications back to the publisher.

Ease of installation, deployment, and use

SQL Server 2000 includes a set of administrative and development tools that improve upon the process of installing, deploying, managing, and using SQL Server across several sites. SQL Server 2000 also supports a standards-based programming model integrated with the Windows DNA, making the use of SQL Server databases and data warehouses a seamless part of building powerful and scalable systems. These features allow you to rapidly deliver SQL Server applications that customers can implement with a minimum of installation and administrative overhead.

Data warehousing

SQL Server 2000 includes tools for extracting and analyzing summary data for online analytical processing. SQL Server also includes tools for visually designing databases and analyzing data using English-based questions.
6.0 EXPERIMENTAL RESULTS AND DISCUSSION

The result analysis allows data to be examined at a level of abstraction appropriate for the programming model of the experimental application. Result analysis is the provision of objective feedback to performers trying to get a positive change in performance. There are three basic steps in the performance analysis process: data collection, data transformation, and data visualization. Data collection is the process by which data about program performance are obtained from an executing program.

Data are normally collected in a file, either during or after execution, although in some situations it may be presented to the user in real time. Three basic data collection techniques can be distinguished: Profiles record the amount of time spent in different parts of a program. This information though minimal is often invaluable for highlighting performance problems.

Outlines typically are gathered automatically. Counters record either frequencies of events or cumulative times. The insertion of counters may require some programmer intervention. Event traces record each occurrence of various specified events, thus typically producing a large amount of data. Traces can be produced either automatically or with programmer intervention. The Table 6.1 represents experimental result for existing Key-Policy ABE Algorithm model. The table shows the selecting the number of Attributes, Access control count and Access permission count. The table contains the various attributes, owner and access policy control and access permission count.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Attribute</th>
<th>User</th>
<th>User Group</th>
<th>Access Control Count [N]</th>
<th>Access Permission Count [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name</td>
<td>A</td>
<td>G1</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>Age</td>
<td>B</td>
<td>G2</td>
<td>6</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>DOB</td>
<td>C</td>
<td>G3</td>
<td>10</td>
<td>200</td>
</tr>
<tr>
<td>4</td>
<td>Salary</td>
<td>D</td>
<td>G4</td>
<td>12</td>
<td>240</td>
</tr>
<tr>
<td>5</td>
<td>Attendance</td>
<td>E</td>
<td>G5</td>
<td>17</td>
<td>340</td>
</tr>
</tbody>
</table>

Table 6.1 Security Analysis of Existing Model (KP-ABE)

The Figure 6.1 represents experimental result for existing Key-Policy ABE Algorithm model. It shows the selecting the number of Attributes and Access permission count based on the access control policy count.
Security Analysis of Existing Model (KP-ABE)

![Graph](image)

**Figure 6.1 Security Analysis of KP-ABE**

The Table 5.2 represents experimental result for the proposed Cipher Text-Policy Attribute-Based Encryption (CP-ABE) with User Revocation model. The table shows the selecting the number of Attributes, Access control count and Access permission count. The table contains the various attributes, owner, access policy control and access permission count.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Attribute</th>
<th>User</th>
<th>User Group</th>
<th>Access Control [N]</th>
<th>Access Permission Count [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name</td>
<td>A</td>
<td>G1</td>
<td>8</td>
<td>160</td>
</tr>
<tr>
<td>2</td>
<td>Age</td>
<td>B</td>
<td>G2</td>
<td>13</td>
<td>260</td>
</tr>
<tr>
<td>3</td>
<td>DOB</td>
<td>C</td>
<td>G3</td>
<td>16</td>
<td>320</td>
</tr>
<tr>
<td>4</td>
<td>Salary</td>
<td>D</td>
<td>G4</td>
<td>22</td>
<td>440</td>
</tr>
<tr>
<td>5</td>
<td>Attendance</td>
<td>E</td>
<td>G5</td>
<td>26</td>
<td>520</td>
</tr>
</tbody>
</table>

**Table 6.2 Security Analysis of proposed CP-ABE**

The Figure 6.2 represents experimental result for proposed Cipher Text Policy based ABE Algorithm model. It shows the selecting the number of Attributes and Access permission count based on the access control policy count.
Figure 6.2 Security Analysis of proposed CP-ABE

The Figure 6.3 represents comparison result of existing Key-Policy ABE Algorithm and Cypher text Policy base ABE. It shows the selecting the number of Attributes and Access permission count based on the access control policy count.

Figure 6.3 Comparisons of Security Properties of KP-ABE & CP-ABE with User Revocation
7.0 Conclusion

The proposed ciphertext-policy attribute-based encryption with user revocation scheme provides a big advantage by supporting user-defined time-specific authorization and fine-grained access control and data secure self-destruction. The proposed scheme allows a data owner to define the access control policy and enforce it on his outsourced data. It also features a mechanism that enables more fine-grained access control with efficient attribute and user revocation capability. It is sent that the proposed scheme is efficient and scalable to securely manage the outsourced data.

The proposed ciphertext-policy attribute-based encryption model does includes the set of the attributes, tree access policy, and the definition of the time instant, because their costs are negligible if compared with the key generation. In this research proposed a novel approach which is Dekey, a new construction in which users do not need to manage any keys on their own but instead securely distribute the convergent key shares across multiple servers.

Security analysis demonstrates that Dekey is secure in terms of the definitions specified in the proposed security model. In addition, the users can revoke from the given group at any time. To do so, if the original (first) user of the group intimates the server with a user’s (B) revocation, then the server rejects the proof of ownership submitted by that user (B). Likewise, session based deduplication is considered.

7.1 Future Enhancements

The data owner encrypts the data to share with users in the system, in which every user’s key is associated with an access tree and each leaf node is associated with a time instant, for this in further it can be enhanced with the logic minimum spanning tree.

Further for the purpose of encryption and decryption process for the user data, Triple DES algorithm can be implemented for providing the better security for the user confidential data.

In future these options can be included so that Session based outsource data access can be provided to increase the security. User Revocation management can also implemented. Key Storage cost can be reduced when compared to existing system.

The following enhancements are should be in future to enhance the performance of the proposed system:

- In future Minimum Spanning Tree logic may applied for the access control tree
- Further the T-DES algorithm can be used for the encryption and decryption process
- Design dynamic mapping functions among keys in the hierarchy and index numbers of data blocks so that it can progressively reorganize the data blocks based on their access patterns for key management.

8.0 References


