A Privacy-Protection Data Processing Solution Based on Cloud Computing

FURUKAWA Jun, FURUKAWA Ryo, MORI Takuya, MORI Kengo, ISSHIKI Toshiyuki, ARAKI Toshinori

Abstract

Following the diffusion of cloud-based services and the consequent increase in the opportunities for handling sensitive data, concerns about leakage and abuse of such data has been increasing. In order to deal with these concerns, this paper introduces suitable countermeasures. These are a technology for preventing data leakage by processing data in an encrypted form and one for protecting data by selecting optimum processing measures according to content. Both of these technologies are results achieved from the "R&D of Security Technology for Promoting Transition to Cloud Services in Preparation for Disasters." This research project was commissioned by the Japanese Ministry of Internal Affairs and Communications (MIC).

Keywords

privacy protection, cloud computing, processing data in an encrypted form, importance level , policy arbitration method

1. Introduction

The diffusion of cloud services is increasing the demand for cloud services for handling sensitive data such as personal data, but there are still many sets of circumstances in which the cloud is not used due to security concerns. In fact, it is very hard for the user to review cloud services directly so that concerns about data leakage and abuse cannot be solved easily. These concerns may be solved using the available data leakage prevention technology by processing encrypted data and the one for protecting data by selecting optimum processing according to content. By solving concerns about leakage of sensitive data it is expected that more services will be enabled to use cloud systems.

The above technologies process the encrypted data while sensitive data is not decrypted in the cloud system. The key for decrypting the encrypted text is therefore in the hands of the data owner, and not in the cloud system where the processing is executed. This means that, even if the cloud service leaks the data, it is the encrypted data that is leaked so that leakage of the actual data can be prevented.

The latter technology that protects data according to its contents processes the data provision destination and applies privacy protection based on the importance level labeled by analyzing the content of the data transmitted from each individual and according to the information protection and usage requirements determined in the "policy." This means that services that utilize personal data while protecting privacy can be implemented safely.

The fundamental idea underlying these solutions is to make attackers unable to extract substantial secret information such as personal data, etc. even if they succeed in obtaining it from the cloud (**Fig. 1**). In 2012, we conducted demonstration ex-

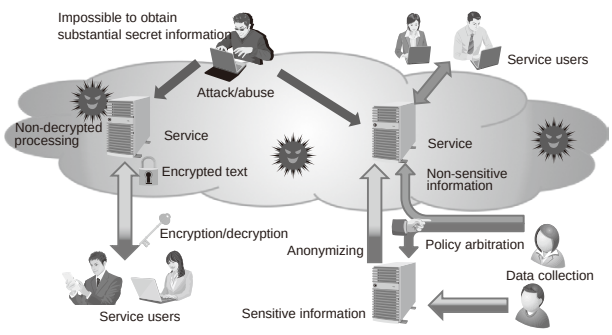

Fig. 1 Processing of unencrypted data and information protection usage policy arbitration method.

periments of these technologies within the framework of the "R&D of Security Technology for Promoting Transition to Cloud Services in Preparation for Disasters," research commissioned by the Japanese Ministry MIC. In this paper, we describe these technologies together with associated matters and report on details of the commissioned research.

2. Processing Non-decrypted Data

In addition to the technology for processing non-decrypted data, there are other technologies that can prevent virtual data leakage by processing encrypted data in the cloud system, which include simple encryption, functional encryption or secret sharing, secure multi-party computation, and fully homomorphic encryption. The following paragraphs describe the characteristics of these technologies.

(1) Simple encryption

If data is encrypted before storage in the cloud system, access to the data can be restricted to the owners of the secret key. However, if there are multiple users, it is necessary to prepare a separate means for distributing the secret key according to their authorizations.

(2) Functional encryption

This encryption method enables implementation of highly accurate access management of multiple users and is convenient for the encryption of shared storage. When this method generates encrypted text, it defines the users who can decrypt the text based on the conditions met by the authorizations of such users. It is for example possible to specify "a job position of X or higher that belongs to group A and does not belong to group D."

The functional encryption controls the user access by providing users with secret keys according to their authorizations. As a result, decryption is possible only by the users owning the secret keys that meet the specified conditions, and is completely impossible in the cloud. Nevertheless, this mechanism makes it necessary to strip the users of their secret keys or rewrite the encrypted text in order to change the authorizations. The former task is difficult to achieve and the latter requires a huge volume of processing.

(3) Secret sharing

This technology saves data by distributing it in several clouds so that the confidentiality of the original data can be protected even if data is leaked from any of the clouds. This method converts data into multiple shares, with which original data cannot be restored unless more than a certain number of shares are collected. The shares are distributed in several clouds to secure confidentiality. Each cloud controls user access to the shares it holds so that this technology is capable of highly accurate access control of the original data via easy authorization changes, which is not possible with the functional encryption.

The method based on secret sharing can secure the availability of data because, even if some clouds lose shares, the original data can be restored from the shares held in other clouds. In addition, this solution also accepts the addition of a forgery detection mechanism for data integrity

The processing operations possible with simple encryption, functional encryption and secret sharing technologies that are outlined above are limited to the read and write operations.

(4) Secure multi-party computation

This technology is capable of arbitrary data processing by overcoming the limitations of secret sharing, which is restricted to read and write processing. After the data has been distributed in several clouds by secret sharing, the clouds can generate the results of arbitrary computations of the original data in cooperation, without restoring the distributed secret data.

However, the processing of secure computations is usually slow and, if users can access the original data, it is much simpler if they read the data and perform the computations themselves. Consequently, this technology is effective for a service that discloses to the users only the simple statistical values obtained by comparable computations of the original system data. The original data cannot be disclosed to either the cloud or the user.

(5) Fully homomorphic encryption

This technology makes it possible to execute arbitrary processing of data in the encrypted status in the cloud, without disclosing data to the cloud. Unlike the secure multi-party computation, the processing is possible in a single cloud. However, this technology also poses problems, which are the very slow processing speed and its drawback of results decoding being restricted to persons who own the secret key.

(6) Processing Non-decrypted data

This technology processes data via a single cloud as with the fully-homomorphic encryption. Although the available processing is limited, it is faster than the fully-homomorphic encryption. The applications depend on the kinds of available processing selected, but they include biometric authentication, analogous chemical compound searches, statistical computations (of mean values, variances and covariances), keyword search, and relational database management system. Decryption of the results is limited to persons who own the secret key in a similar manner to the fully-homomorphic encryption. However, it can sometimes be made available for multiple users by using a technology called proxy re-encryption. This technology enables practical service deployment while providing strong information leak resistance via encryption. This is one of the recommended privacy-protection services that

are introduced in this paper that are subjected to demonstration experiments in the research commissioned by the Japanese MIC. It also belongs to the services based on non-encrypted processing.

3. Policy Arbitration Method Technology

3.1 Policy Approach

The access control or privacy policies endorse the affirmation or rejection of an operation based on the shared data, operational obligations, and suitable conditions for permitting the operation, etc. With regard to privacy protection, the privacy policies define the affirmation/rejection of data acquisition and the processing of privacy protection is enforced when the relevant data is made available, etc.

In the present research, we developed an architecture in which the front end of the cloud makes a decision and enforces privacy protection automatically, and optimum privacy protection is achieved regardless of the method of handling the data of the various services in the cloud (**Fig. 2**).

In this architecture, the control technology of the virtual server, which makes a decision and enforces the policy, enforces optimum data protection processing based on the data protection policy generated by the data protection policy generation function, the importance level of data, and the security status of the service.

Below, we detail the policy arbitration technology used by the data protection policy generation function.

3.2 Policy Arbitration Method Technology

The policy arbitration method technology generates optimum data protection measures according to the security status and data importance level by reducing the burden on users and services.

The policy arbitration method is accompanied by the issues defined below. The methods of solving them will be described separately in the following subsections.

1) **Conflict between user and service policies**
Difficulty of solving conflict between the data protection requirements of the user and the data usage requirements of the service.
2) **Compatibility with various service environments**
Difficulty of generating a data protection policy that is optimized for the security status and the data importance level.

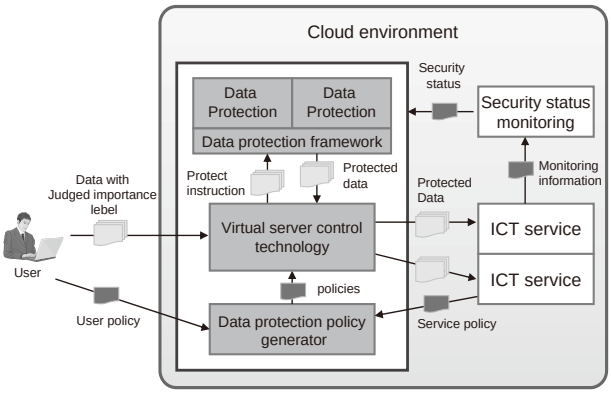

Fig. 2 Cloud service framework using a user policy method.

3.2.1 Arbitration Method Technology Based on Policy Ranking

In order to resolve the conflict between policies, it is necessary to create acceptable policies based on communications between the user and the service providers in order to reach a compromise regarding their needs. However, the need for a large number of communications between the user and service provider poses the problem that a heavy burden is placed on both parties.

We have therefore developed an arbitration method that utilizes the policy ranking for reducing the volume of communications. With this method, the service provider registers acceptable policies in advance and presents service policies by ranking them according to that input by the user. The user can select an acceptable policy by referencing the presented policy rankings.

We propose to allow the two parties to find the procedure to agree on a policy that is acceptable to both of them via conducting minimal communications. Evaluation revealed that the time required for the speed to present the policy ranking is as high as about 500 msec., when the number of service procedures is around 200.

3.2.2 Arbitration Method Based on Inter-user Similarity

In order to protect data optimally according to the environments characterized by the security status and data importance level it is necessary to use a procedure that matches each environment. Nevertheless, it is difficult for the user to set an acceptable policy for all of the environments.

To solve this problem, we developed an arbitration technology that recommends the optimum data protection policy to suit the current environment. This technology computes the similarity levels of users based on the policies that have been generated for the various environments and recommends user policies that show high similarity.

This solution makes it easy to generate policies to suit various environments by recommending optimum policy procedures. The evaluation showed that the time required for the similarity level evaluation speed is as high as about 500 msec. when the number of users is 1,000.

4. Demonstration Experiments

4.1 Privacy-protection Recommendation Processing

In the event of a large-scale disaster, it becomes important to utilize private rented accommodation effectively as the hasty preparation of a large number of temporary housings is a difficult task. Since direct real estate mediation by administrators involves various issues, it is desirable to utilize general rental agreements. However, the time taken for the screening of borrowers by the lenders becomes a significant obstacle to quick conclusion of general rental agreements in the event of a large-scale disaster.

We suggest speeding up improvements to the residential environment in order to reduce the number of applicant screenings and to recommend housings for the disaster sufferers that would not be refused.

Screening requires personal information including the economic circumstances of the borrowers and their guarantors. This means that personal information is also required for housing recommendations that would not be refused in screenings. This leads to concerns regarding possible information leakage when the service is performed in the cloud. We have therefore solved such information leakage concerns by selecting the mechanism for processing information without decrypting from the encrypted status. We have also created a rental housing mediation service based on this mechanism and have conducted demonstration experiments within the framework of the research commissioned by the Japanese MIC. Here, the processing required for the service is implemented by computing the inner product of vectors in the encrypted state. We selected the expression by vectors that express the circumstances of the borrower and the conditions of the lender in vectors and computed the degree of their matching as the inner product.

We obtained certain evaluations from the demonstration experiment because it presents a potential for implementing new services by applying the present technology for processing information in the un-decrypted state. The participants in the experiment proposed ideas for services using this technology, which suggested the potential of wide applications for the technology.

4.2 Cloud-based Information Exchange Service

If a large-scale disaster occurs, it is desirable that the local governments and sufferers share efficiently the information useful for enforcement and promotion of anti-disaster activities.

With the present study, we established the following hypotheses and built an information collection/transmission system based on cloud service by developing the policy-based framework described in section 3.

1) To build an information collection/transmission service in the cloud would enable disaster resistant information collection.
2) Allocating each piece of provided information to the optimum services would enable efficient information collection by transmitting various kinds of information to the relevant local governments, residents and governmental institutions.
3) Automatic execution of privacy protection according to the policy of each disaster sufferer would reduce the number of privacy-related concerns and promote the active provision of information.

We verified the hypotheses above in the demonstration experiments for the research commissioned by the Japanese MIC. As a result, 97% of the victims agreed that such a service is useful for maintaining business continuity; almost 100% agreed that the information collection/arrangement would be actually useful in the event of a disaster, and 83% agreed that privacy protection promotes the provision of information. This demonstration experiment allowed us to verify the hypotheses above and conclude that the cloud-based information exchange service would be useful in disaster events.

5. Future Deployment

By applying the encrypted information processing technology, we aim to implement effective services, including the provision of relational databases in the encrypted state or of biometric authentication using encrypted data. By introducing a secure policy based privacy protection technology, we will be able to implement services for utilizing highly sensitive big data such as personal data, by applying anonymization.

Authors' Profiles

FURUKAWA Jun
Principal Researcher
Cloud System Research Laboratories

FURUKAWA Ryo
Cloud System Research Laboratories

MORI Takuya
Principal Researcher
Cloud System Research Laboratories

MORI Kengo
Assistant Manager
Cloud System Research Laboratories

ISSHIKI Toshiyuki
Assistant Manager
Cloud System Research Laboratories

ARAKI Toshinori
Assistant Manager
Cloud System Research Laboratories

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.8 No.1 Solving Social Issues Through Business Activities

Vol.8 No.1

September, 2013

Special Issue TOP

Remarks for Special Issue on Solving Social Issues Through Business Activities

The Reinvention of NEC as a" Social Value Innovator" - Contributing to solving social issues through business activities -

◇ Papers for Special Issue

Build reliable information and communications infrastructure

Features of the Next-Generation Traffic Control System as Seen in an Introductory Example at the Shin-Tomei Expressway

Enabling International Communications
- Technologies for Capacity Increase and Reliability Improvement in Submarine Cable Networks

Component Technologies and Packet-Optical Integrated Transport Systems to Support Core Networks

Development of Technology to Control Radio Signal Interference for LTE Femtocell Base Stations to Achieve Stable Communications Quality Anywhere

Address climate change and environmental preservation

Regular Observation by Global Change Observation Mission 1st-Water GCOM-W1(SHIZUKU)

Express5800 Server Series and iStorage M Series Storages Contributing to Data Center Power Saving

Possibilities in Thermoelectric Conversion Using a New Principle:" Spin Seebeck Effect"

Establish a safe and secure society

CONNEXIVE Ionizing Radiation Measurement Solution

Disaster Prevention Administrative Radio System in Municipality (Broadcast via PA Systems)
- Achievement of Greater Diversity in Disaster Information Transmissions

Promoting the Digitization of Japanese Fire Prevention/Emergency Wireless Communications Systems

NEC's BC Solutions: HYDRAstor - Supporting Business Continuity of Enterprises

Underwater Surveillance System to Counteract Associated Underwater Threats

A Surveillance System Using Small Unmanned Aerial Vehicle (UAV) Related Technologies

A Privacy-Protection Data Processing Solution Based on Cloud Computing

Towards Trustworthy Cloud Storage

Include everyone in the digital society

A Solution to Prevent Wandering by Geriatric Patients - A Validation Test to Ensure Safety in Nursing Care Facilities

Remote Summary Transcription System for the Hearing Impaired

Communication Activation Technology for Suggesting Conversational Topics