ATLANTIS PRESS
Advances in Computer Science Research, volume 50
4th International Conference on Electrical & Electronics Engineering and Computer Science (ICEEECS 2016)

Research on the Data Fusion Method for Smart Grid Cloud Storage

Shaomin Zhang[a], Qing Zhao[b], Baoyi Wang[c]

School of Control and Computer Engineering, North China Electric Power University, Baoding, 071003, China

[a]email:zhangshaomin@126.com, [b]email:1349878647@qq.com, [c]email:wangbaoyiqj@126.com

Keywords: data fusion; updated key; cloud storage; dynamic broadcast encryption

Abstract. For the question that it cannot ensure the privacy of identity and real-time updated key in the process of data fusion in smart grid, this paper proposes a algorithm of data fusion based on dynamic broadcast encryption. The control center broadcasts the key through dynamic broadcast encryption. Users encrypt and sign the data, then upload to the cloud storage. The cloud storage fuses the data and signatures. Then the control center can analyze the data which stored in the cloud storage. The algorithm in this paper supports the privacy protection of users and real-time updated key. Finally the algorithm is proved to be correct and effective by the theory analysis and experiment.

Introduction

The concept of smart grid is to get more information about how users use electricity, then make the generation, distribution and consumption of electricity more optimize[1].The smart grid uses the modern network, communication and information technology to complete interaction of huge amounts of data, and realizes exchange and sharing of data among the equipments of smart grid [2].Large amounts of data in smart grid need real-time communication, so it makes a lot of requests about the computing power and the capabilities of network communication in system [3]. Data fusion is a technology of automatic information processing, it plays a very important role in smart grid., It can reduce the amount of network data transmission, and it can reduce network congestion to send, improve the performance of network in the transmission of sensor data. In the center of the data it can make raw data to be understanded information or decision[4]. So, it is very important to fuse the data between different systems and equipments through the cloud storage [5-7].

Literature [8] proposed a fusion method which can provide the protection of data integrity, introduced a signature of end-to-end and generated homomorphic signature for the results. However, all of the electric meter uses the same key in fusion, it is vulnerable to attack. Literature [9-10] opposed a fusion method which is safety and effective, it completed the multidimensional data fusion through the use of homomorphic Paillier password system. But the session key is fixed, it can not guarantee the security of forward.

This paper proposes a algorithm of data fusion based on cloud storage. It uses an open channel to broadcast keys by dynamic broadcast encryption, manages users to add or revoke effectively, and guarantee the identity of the user privacy and the security.

Data Fusion Model

Smart meters is the terminal equipment of smart grid, they have the function of storing users information about themselves and electricity, and they have multiple transmission methods of data, it called two-way communication and other functions[11-12],This article is mainly aimed at fusing data in smart meters of users. The model of data fusion is shown in figure 1,it mainly divided into three parts, control center, cloud storage and users. Where the control center is a credible institution to be responsible for managing initialization of system, generating parameters of system, managing the operate of adding and revoking users; The cloud storage is responsible for fusing the cipher tests of users, and integrating the information, signature and timestamps, then sending them to control center; N users is data source that they provide the actual data of power consumption and demanded

data of all users[13].

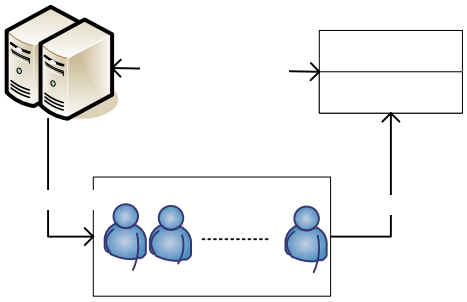

Fig.1.Model of data fusion

To complete the data fusion of n users, each user need to divide and signature the data. Then users send the data and signature to the cloud storage; After receiving the message from users, the cloud storage fuses and stores the data and signature; The control center can be verified to determine whether the data in the cloud storage is stored rightly. The control center can decrypt and analysise the data in cloud storage after verification of data, thus it can improve the safety and save resources of storage and transmission.

Data Fusion Algorithm

The algorithm of data fusion is mainly divided into four stages, generation of key, initialization, data fusion and verification of integrity. First of all, the control center generate key, then broadcast key to every user through DBE. Users use received key to block the file and sign, then send messages to the cloud storage. The cloud storage store and fuse messages. The data stored in the cloud storage can be verified by the control center, and it can ensure the data is stored rightly. The control center can decrypt and analyse the messages after downloading from the cloud storage.

1)generation of key

The control center randomly generated the private key$\mathrm{SK}=\epsilon_0 \overset{R}{\leftarrow} Z_q^*$,$\mathrm{v}=g^{\alpha\epsilon_0}$,$k_0=g^{\epsilon_0}$.then chosed $\alpha \overset{R}{\leftarrow} Z_q^*$ randomly, and generated$\{g^{\alpha^j}\}, 0\le \mathrm{j}\le \mathrm{k}+1$.

Primary key$\mathrm{PK}=\{\mathrm{g},\mathrm{u},\mathrm{q},\mathrm{v},\{g^{\alpha^j}\}0\le \mathrm{j}\le \mathrm{k}+1,k_0\}$, Public key$\mathrm{MK}=\{\epsilon_0,\alpha\}$,Secret key$\mathrm{SK}=\epsilon_0 \overset{R}{\leftarrow} Z_q^*$.

The group of dynamic broadcast is U$\{\mathrm{U}|U_i\in \mathrm{U},1\le \mathrm{i}\le \mathrm{d}\}$.The article uses DBE to encrypt the secret key $\mathrm{DBE.Enc}_{ek}(\mathrm{SK},\mathrm{U})$,where ek is encrypted key of group. The mainly user broadcasts the dycrypted key to every use through channel, then the user U_i calculate SK through $\mathrm{SK}=\mathrm{DBE.Dec}_{dki}(DBE.Enc_{ek}(SK,U))$,where dki is the dycrypted key of user U_i.

Including, p and g is prime order of G_1 and G_2.It contains bilinear map$\mathrm{e}: G_1\times G_2\to G_2$,The parameters is $(G_1,G_2,\mathrm{p},\mathrm{e},\mathrm{g},\mathrm{H},\eta_1,\cdots,\eta_k)$,where H is hash function $\mathrm{H}:\{0,1\}^*\to G_1,(\eta_1\cdots\eta_k)\in G_1$.Each block has k elements, the sum is n. As follows: $\mathrm{M}=(m_1,\cdots m_n),m_i=(m_{i,1},\cdots m_{i,k})\in Z_p^k$.The number of group is d, the original user generates the corresponding parameters for DBE.

2)Initialization

Each user divides the file $\mathrm{F}_t,1<t<\mathrm{U}$ into n blocks, each block is divided into k elements. As follows: $m_i=(m_{i,1},\cdots m_{i,k})\in Z_p^k$.The way to calculate the signature $\sigma_i,1\le i\le n$ of blocks m_i is

$$\sigma_i=(u^{B_i}\cdot\prod_{j=0}^{k-1}g^{m_{i,j}\alpha^{j+2}})^{\epsilon_0}=(u^{B_i}\cdot g^{\frac{f_{\rightarrow}(\alpha)}{\beta_i}})^{\epsilon_0} \quad (1)$$

where $\underset{\beta_i}{\rightarrow}=(0,0,\beta_{i,0},\beta_{i,1},\cdots,\beta_{i,k-1}),\beta_{i,j}=m_{i,j}\cdot B_i=H(\{f_{name}||i||t_i\})$.The name of file is f_{name},the index of block m_i is i, and the timestamp is t_i.

3)data fusion

Users send the files and corresponding signatures to the cloud storage. Then the cloud storage fused the message, and get the signature $\sigma = \sum_{i=1}^{n} \sigma_i$ and the file $F = \sum_{t=1}^{U} F_t$.

4)verification of integrity

When the users in the group receive the private key, they all have right to modify the data.

The data stored in the cloud storage can be verified by the control center, and it can ensure the data is stored rightly. The process of verification is as follows:

The control center randomly chooses d blocks as data set D, then generates two random member R and μ,$X = \{(g^{\varepsilon_0})^R\}$.The information of challenge is $CM = \{D, X, g^R, \mu\}$.When the cloud storage receives CM, it calculates the message

$$y = f_{\vec{A}}(\mu) \bmod q \tag{2}$$

where $\vec{A} = \{0, 0, \sum_{i \in D} p_i m_{i,k-1}\}, \{p_i = \mu^i \bmod q\}, i \in D$.Then the cloud storage calculates the message

$$\varphi = \prod_{j=0}^{k} (g^{a^j})^{w_j} = g^{f_{\vec{w}}(a)} \tag{3}$$

where $\vec{w} = (w_0, w_1, \cdots, w_k)$,and

$$\pi_i = e(\sigma_i, g^R) = e(\left(u^{B_i} \cdot g^{f_{\vec{\beta}_i}(a)}\right), g)^{\varepsilon_0 R} \tag{4}$$

where $\pi = \sum_{i \in D} \pi_i^{p_i}$.Finally the cloud storage generate the message of proof $(Prf = \{\pi, \varphi, y\})$.

The control center calculates $\eta = u^w$,where $w = \sum_{i \in D} B_i p_i$.Then it need to verify the equation as follows:

$$e(\eta, k_0^R) \cdot e(\varphi^R, v, k_0^{-\mu}) = \pi \cdot e(k_0^{-y}, g^R) \tag{5}$$

If the equation is set up, then returns 1.It represents that the result of integrity verification is success. Otherwise, it returns 0. It represents that the result of integrity verification is fail.

Theoretical Analysis

Analysis of data integrity. The premise of this article is that the control center is safety. In this premise, there are two possible that the users will be tampered or deleted: 1) The data will be tampered or deleted during transmission.2) The data will be tampered or deleted in the process of data fusion and stored in the cloud storage.

For the first question, all data are encryptied by homomorphic function, namely it is that all data is a cipher in the process of transmission in the network. It means that only users and the trusted control center can decrypt and get the cipher. And the other attacker who do not have the secret key of data unable to get data. This can ensure the security of data in the process of transmission.

For the second question, the data that stored in the cloud storage can determine the security of data by verification of integrity , which verify the correctness of the algorithm proposed in this paper, namely it is to verify the equation (5) is established.

$$\begin{aligned} \pi \cdot e\left(k_0^{-y}, g^R\right) &= \prod_{i \in D} e\left(\left(u^{B_i} \cdot g^{f_{\vec{\beta}_i}(a)}\right), g\right)^{R\varepsilon_0} \cdot e(g^{-y}, g)^{R\varepsilon_0} = \prod_{i \in D} e(u^{B_i}, g)^{R\varepsilon_0} \cdot \prod_{i \in D} e(g^{f_{\vec{\beta}_i}(a)}, g)^{R\varepsilon_0} \cdot e(g^{-y}, g)^{R\varepsilon_0} \\ &= \prod_{i \in D} e(u^{B_i}, g)^{R\varepsilon_0} \cdot e(g^{f_{\vec{A}}(a)}, g)^{R\varepsilon_0} \cdot e(g^{-y}, g)^{R\varepsilon_0} = e(u^{\sum_{i \in D} B_i}, g)^{R\varepsilon_0} \cdot e(g^{f_{\vec{\omega}}(a)}, g^{(a-\mu)})^{R\varepsilon_0} \\ &= e(\eta, k_0^R) \cdot e(\varphi^R, v \cdot k_0^{-\mu}) \end{aligned} \tag{6}$$

Privacy protection of users. In the process of traditional integrity verification, the publicly verifier can easily identity the signers of data blocks. Then it can get which data blocks or users is even more important target. If the date be blabbed by the public, it will be very dangerous. So it is necessary to protect the privacy of users.

In this method, the private key is generated by the control center, then encrypted and broadcasted dynamically by safety channel to the users of group U. Thus the users use the same key to encrypt its own data, the public cannot get which data blocks or users is even more important target.

Addition of users. The process of adding or revoking users all need to do operation of update. in order to reduce the computational overhead in the process of key update. this article uses DBE to realize real-time update of key.

The mainly user adds user u' ($u' \in U'$)to the group U, then it gets $U' = \mathrm{U} \cup u', u' \notin U$.The mainly user uses encrypted key to encrypt secret key repeatly $\mathrm{DBE.Enc}_{ek}(\mathrm{SK}, U')$,then broadcasts to the group U'.The mainly user send decrypted key to user u',then the user u' to get the secret key $\mathrm{SK} = \mathrm{DBE.Dec}_{dk'}(DBE.Enc_{ek}(SK, U'))$.Where the decrypted key of user u' is dk'.The added user can use SK to read and write data blocks, also can signature the data blocks.

Revocation of users. If there is a user who wants to be revoked, the original user generates a new secret key $\mathrm{SK}' = \epsilon_0{}'$,and a new public key $PK' = \left\{g', u', q', v', \left\{g^{\alpha^j}\right\} 0 \le j \le t, k_0{}'\right\}$.Where after revoking, the number of the group is t, the group is $u'\mathrm{U} = \mathrm{U}' \cup u', u' \notin U'$.The mainly user encrypts secret key SK' repeatly $\mathrm{DBE.Enc}_{ek}(sk', u')$,then broadcasts it through safety channel. Every user can receive SK', but only the revoked user cannot receive the secret key SK' because of $u' \notin U'$.At the same time, the mainly key repeatly calculates public key $\mathrm{RK} = \frac{\epsilon_0{}'}{\frac{R}{\epsilon_0 + Z_p^*}}$,and send it to the cloud storage. When the cloud storage receives the PK, it will re-sign all data blocks

$$\sigma_i{}' = \sigma_j{}^{rk} = (u^{B_i} \cdot \prod_{j=0}^{s-1} g^{m_{i,j}\alpha^{j+2}})^{\epsilon_0{}'}$$

$$= (u^{B_i} \cdot g^{\frac{f_{\rightarrow}(\alpha)}{\beta_i}})^{\epsilon_0{}'} \tag{7}$$

Experiment And Result Analysis

In this paper, the experimental environment is established by four nodes of Hadoop cloud platform in the laboratory, the machine configuration of node is Intel(R)Core (TM)i5-2400 4-core CPU@2.60 GHz，4 GB RAM, the network bandwidth is 100 Mbit/s, and the version of Hadoop is 0.20.2. The hardware configuration of TPA machine is Linux,3.4GHz Intel i7-3770 CPU,16GB,the hardware configuration of users' machine is Linux,2.50GHz Intel i5-2520M CPU,8GB.

Experiment of integrity. We sign and store three values (A, B, C),where to delete 10% for A, modify 10% for B, and do nothing for C. Finally, we do integrity detection for these values. The results are shown in table 1.

Tab.1. Result of detection

File	Result
A	Fail
B	Fail
C	Success

Through the experiment, we get that the file cannot br success after tampered or deleted, only the files that do not be tampered or deleted can be success.

Experiment of updated key. In this scheme. When a new member join in the group, the original user only need to dynamically broadcast secret key to new members. As shown in table 2,adding a new user need 0.13s;When a user wants to be revoked, the cloud storage re-sign the corresponding data block. Legitimate users do not need to download and sign the data block, this time takes 2.91s.Compared with the literature [14], although the time when adding users need more than 0.02 s. But the literature [14] need users to re-sign blocks when revoking users, this takes a longer time.

Tab.2. Comparision of time between adding users and revoking users

	adding	revoking
This scheme	0.15s	2.93s
Literature [14]	0.12s	43.96s

Experiment of overhead time. With the rapid development of smart grid, the number of users is also rised by straight line. The algorithm in this paper is designed in the basis. When the number of users is different, it guarantees the time of verification remain the same, As shown in figure 2, when the number of users increase, the algorithms in this article remain at about 2.14s.At the same time, it proves that the method of data fusion is effective in this paper.

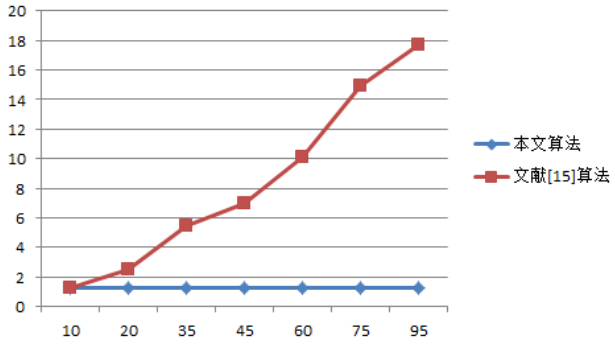

Fig.2. Relationship between number of users and time of verification

Summary

This paper provides a method of data fusion based on the cloud storage. The control center broadcasts the key through DBE. Users encrypt and sign the data, then upload to the cloud storage. The cloud storage fuses the data and signatures, Then the control center can analyze the data which stored in the cloud storage. The algorithm in this paper supports the privacy protection of users and real-time update of key. Finally the algorithm is proved to be correct and effective by the theory analysis and experiment.

Acknowledgement

In this paper, the research was sponsored by the National Natural Science Foundation of China (Project No. 61300040) and Scientific Research Project of Hebei Province (Project No. Z2012077).

References

[1] Zhang Dongxia, Miao Xin, Liu Liping, Zhang Yan, Liu Keyan. Research on Development Strategy for Smart Grid Big Data[J].Proceedings of the CSEE,2015,35(1):2-12.

[2] Peng Xiaosheng, Deng Diyuan, Cheng Shijie, Wen Jinyu, Li Zhaohui, Niu Lin. Key Technologies of Electric Power Big Data and Its Application Prospects in Smart Grid[J]. Proceedings of the CSEE, 2015, 35(3):503-511.

[3] Chen Liang, Lin Yongfeng. Homomorphic encryption based security data fusion technology of smart grid[J].Modern Electronics Technique, 2016,39(9):82-86.

[4] Li Xiangyang, Li Lingjuan,Chen Jianxin, Xu Xiaolong. Research on Application of Data Fusion in Smart Grid[J].Computer Technology and Development, 2012,22(4):215-218

[5] Song Yaqi, Zhou Guoliang, Zhu Yongli. Present Status and Challenges of Big Data Processing in Smart Grid[J].Power System Technology, 2013,04:927-935.

[6] Pan Xinzhi, Zhang Wansheng. Research and Application of Information Fusion Technology in Smart Substations[J].

Telecommunications for Electric Power System, 2013,01:36-41.

[7] Cao Junwei, Wan Yuxin, Tu Guoyu, Zhang Shuqin,Xia Aixuan, Liu Xiaofei,Chen Zhen, Lu Chao. Research Institute of Information Technology[J].Chinese Journal of Computers. 2013, 01:143-167.

[8] Cao Junwei, Wan Yuxin, Tu Guoyu, Zhang Shuqin, Xia Aixuan, Liu Xiaofei, Chen Zhen, Lu Chao. Research Institute of Information Technology[J].Chinese Journal of Computers. 2013, 01:143-167.

[9] Wen M, Lu R, Lei J, et al. ECQ: An Efficient Conjunctive Query scheme over encrypted multidimensional data in smart grid[C]. Global Communications Conference (GLOBECOM), 2013IEEE. IEEE, 2013:796-801.

[10] Li H, Liang X, Lu R, et al. EDR: An efficient demand response scheme for achieving forward secrecy in smart grid[C]. Global Communications Conference (GLOBECOM), 2012 IEEE. IEEE,2012:929-934.

[11]Liu Yan.A Secure Data Aggregation Scheme Based on Homomorphic Encryption in Smart Grid[D].Beijing Institute of Technology,2015

[12]Jiawei Yuan,Shucheng Yu.Efficient Public Integrity Checking for Cloud Data Sharing with Multi-User Modification[J]. IEEE transactions on information forensics and security. 2015: 1556-6013.

[13]Wang Boyang.The Study of Public Auditing for Shared Data in the Cloud[D].Xidian University.2014.

[14]D. Song, E. Shi, I. Fischer, and U. Shankar, Cloud Data Protection for the Masses, IEEE Computer, 2012, 45(1):39–45

[15]J. Yuan and S. Yu, “Proofs of retrievability with public verifiability and constant communication cost in cloud,” in Proceedings of the 2013International Workshop on Security in Cloud Computing, ser. CloudComputing ’13. Hangzhou, China: ACM, 2013, pp. 19–26.