A Privacy Protection Method Based on CP-ABE and KP-ABE for Cloud Computing

Yi-mu Ji [a, *1], Jie Tan [a], Hai Liu [b], Yan-peng Sun [a], Jia-bang Kang [a], Zizhuo Kuang [a], Chuanxin Zhao [c]
[a] College of Computer, Nanjing University of Posts and Telecommunication, China
jim4njupt@gmail.com
[b] School of Computer, South China Normal University, Guangzhou, China
jim_love@163.com
[c] Department of Computer, Anhui Normal University, Wuhu, China
zcxonline@126.com

***Abstract*—Most of the existing ways for strengthening users' confidence in uploading private data to a cloud platform stress too much on security. On the contrary, cloud computing must be open and distributed. Moreover, it should provide highly robust, secure, and quality services to users. Therefore, establishing a balance between the performance and the security of cloud platforms is necessary. In this study, a hybrid privacy protection solution based on key policy-attribute-based encryption (ABE) and cipher policy-ABE is proposed. In this solution, privacy information is encrypted based on user attributes and cloud service type. This novel solution is verified by a case study on a campus cloud environment.**

***Index Terms*—cloud computing, attribute-based encryption, data security, privacy protection**

I. Introduction

Cloud computing helps enterprises in cutting costs for hardware, platform, and software [1, 2]. However, service storage via third-party cloud computing service providers poses security risks, particularly those pertaining to privacy protection [3, 4]. On the one hand, user identity should be verified and validated. On the other hand, user privacy data stored in cloud platforms should be secure to prevent cloud platform service providers or other customers from accessing such data. To avoid damage to and loss of cloud data, cloud service providers must secure and validate the completeness of such data. Currently, privacy data can be protected using several technologies, and many government bodies have formulated and implemented laws to regulate the practices and obligations of cloud service providers.

Technologies for privacy protection have been studied by many scholars. Ref. [5] proposed one public key encryption to authenticate user identity. Ref. [6, 7] proposed a homomorphism encryption to prevent cloud service providers from reading the contents of data when executing data computation for users. A fuzzy algorithm can be used to protect user data [8]; that is, users can code and upload the data to a cloud server. During this process, the end server does not gain access to the real data of the users.

The studies above are clearly more focused on data security than on service performance and user experience. A balance between security and performance must therefore be established. Sahai and Waters proposed the concept of attribute-based encryption (ABE) [9]. ABE allows text and private key to be ciphered based on one group of attributes. Users can then decrypt the cipher text when their private key matches the attributes of the cipher text. In this way, a balance is established between the performance and privacy protection capabilities of the platform. The following two methods based on ABE have been proposed: key policy (KP)-ABE and cipher policy (CP)-ABE. These two methods can classify the attributes of user data. An access control tree (ACT) is created based on the classification levels [10]. In the following section, a solution that combines KP-ABE and CP-ABE is proposed.

II. Related Works

KP-ABE was first proposed by Goyal et al. [11], and CP-ABE was proposed based on both KP-ABE and ABE [12]. KP-ABE ensures correlation among the same attribution sets, and ABE can encrypt the attribution set. However, ABE is not capable of identifying the users who might access the encrypted data files. In KP-ABE, the key attributes of users are connected to an ACT comprising the same attributes. Users cannot access and read the data files by relying on a third-party key authority. The rights of users cannot be decided by an encrypted actor. In CP-ABE, the data owner can encrypt the data by creating a corresponding ACT and identifying the users who can access the cipher text. Therefore, the right to access and control data lies on the data owner. The process of CP-ABE is shown in Fig. 1.

This work was supported by the Projects of Jiangsu Industry (BE2010057), Jiangsu Natural Science Foundation (BK20130876), and National Post-doctoral Foundation(2013M541702).

* Corresponding author

doi:10.4304/jsw.9.6.1367-1375

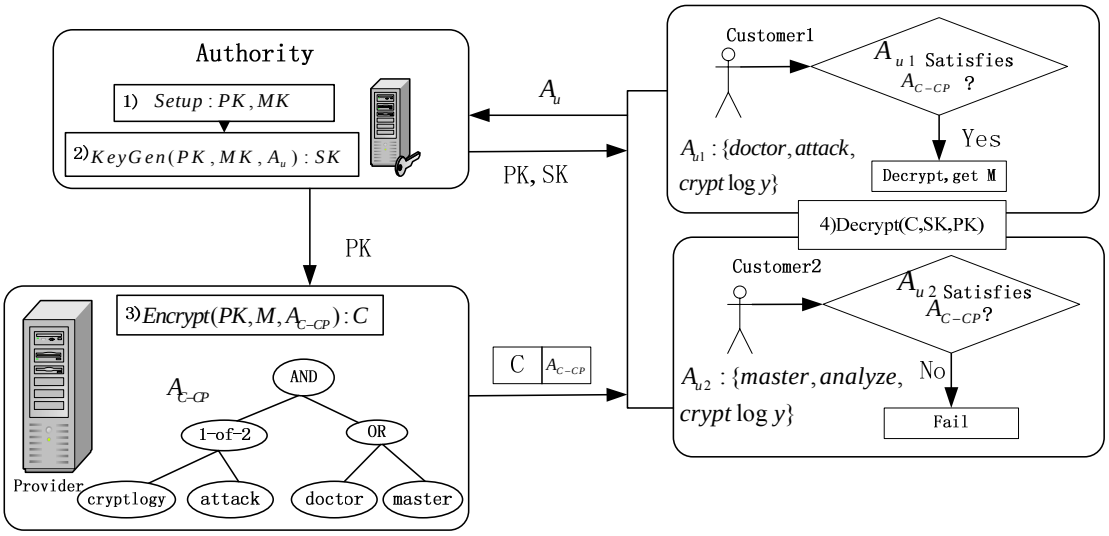


Figure 1. Encryption and decryption process of CP-ABE.

In Fig. 1, A_{C-CP} denotes the access policy of the cipher text and is represented by a tree structure. A_{u_i} is the set of attributes i of the customer's privacy data. PK, MK, and SK denote the private key, main key, and session key, respectively. C denotes the cipher text. The encryption and decryption process of CP-ABE involves four steps: <Set up>, <Key Gen>, <Encrypt>, and <Decrypt>. Meanwhile, the CP-ABE system plays three roles: customer, provider, and authority. The details of each phase in Fig. 1 are introduced below.

Phase 1 <Set up>: In this phase, PK and MK are created by the authority after receiving the request from the customer, who submits A_{u_i} to the authority.

Phase 2 <Key Gen>: The authority generates the SK according to A_u, PK, and MK and sends the PK and SK to the customer.

Phase 3 <Encrypt>: The provider uses PK and A_{C-CP} to encrypt the customer data and generates C. This phase is given by $Encrypt(PK, M, A_{C-CP}): C$. Then, the provider sends C and A_{C-CP} to the customer.

Phase 4 <Decrypt>: Upon receipt of C and A_{C-CP}, the customer determines whether A_u satisfies the policy of A_{C-CP}. If the condition is satisfied, the customer decrypts C. Otherwise, the customer cannot decrypt C.

To flexibly describe the protection policy of the cipher text and to express the complex logical relationship between CP-ABE and KP-ABE, both methods are integrated into the proposed solution. With this novel solution, cloud data can be shared flexibly, and cloud platforms become highly flexible, reliable, and usable.

III. Relationship of User Attributes and Access Control Policies

User privacy data stored in cloud can be classified into different levels according to different security requirements. Therefore, the access control policy (ACP) based on attributes is designed according to different levels. Privacy data are encrypted and stored, thereby allowing users to access privacy data at different levels. Other data are either shared with customers who have rights or owned solely by the data owner. This individual classification according to different security levels is thus flexible and applicable. The mapping of the ACP and the levels of user privacy data is shown in Fig. 2.

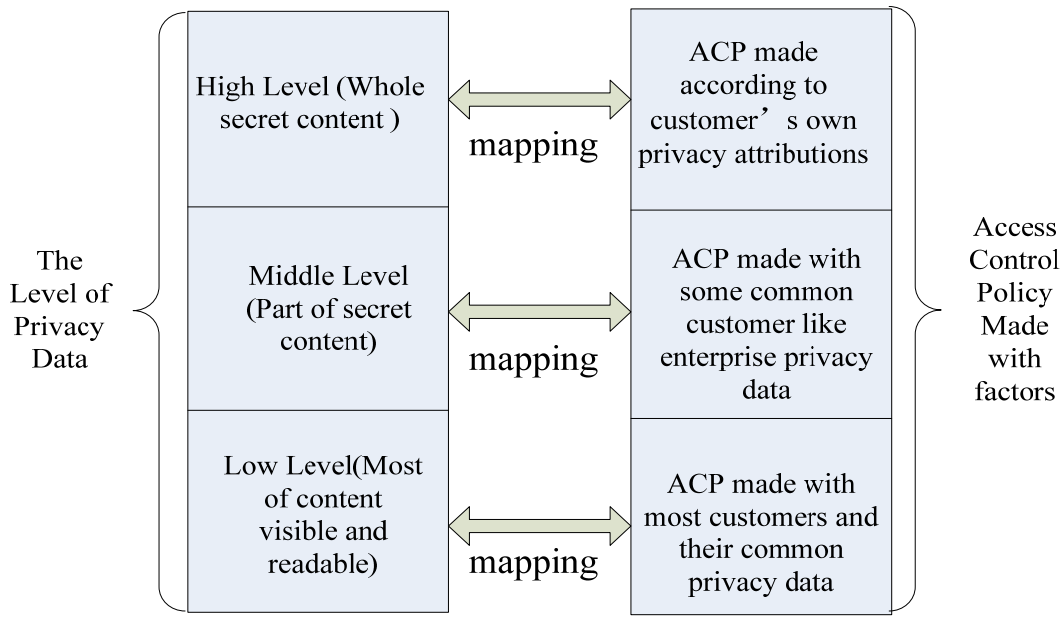


Figure 2. Mapping between ACP and the level of customer privacy data.

The construction of an ACT (Fig. 2), in which the leaf nodes are the privacy attributes, is particularly challenging. The parent nodes are the logical operators, namely, "AND" and "OR." The ACT can encrypt sensitive data, which can then be decrypted when the user attributes satisfy the ACP. An ACT (Fig. 2) comes in three types: high secret level, middle secret level, and low secret level.

- **High secret level**

User data with high secret level holds the most sensitive and private information in a cloud environment. Examples of such information include user mailbox ID and key, address, and bank card ID and key. These data are not known to others. If a user satisfies the ACT, as shown in Fig. 3, the sensitive and private data can be encrypted and read. In Fig. 3, the user owns the real name, department, and member ID or the real ID and key. The user can also decrypt the secret data. Otherwise, the cloud platform rejects the user requirement. In this case, only the valid user can access the data.

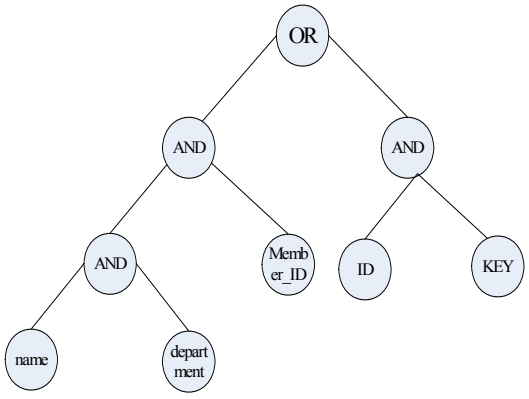


Figure 3. High level ACT structure.

- **Middle secret level**

Some users can access the same data, and they are assumed to be owners of the same right. Take for example an enterprise user. The cloud service provider sets the ACT for the enterprise, as shown in Fig. 4. Fig. 4 indicates that research managers, business sales representatives, and enterprise heads can decrypt encrypted data with middle secret level ACT.

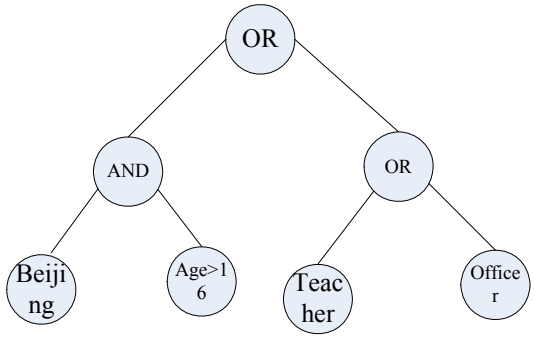


Figure 4. Middle level ACT structure.

- **Low secret level**

If user data, such as books and favorite videos, can be shared with most users, then such data can be shared in a cloud platform and be enhanced in terms of usability. In this case, the service provider can set the ACT such that many users are given the right to read and browse these data. This solution fits different groups. Fig. 5 shows an example of this solution, in which users in Beijing aged over 16 years or teachers or office representatives can access the data. This type of ACT structure applies to data with little sensitive information.

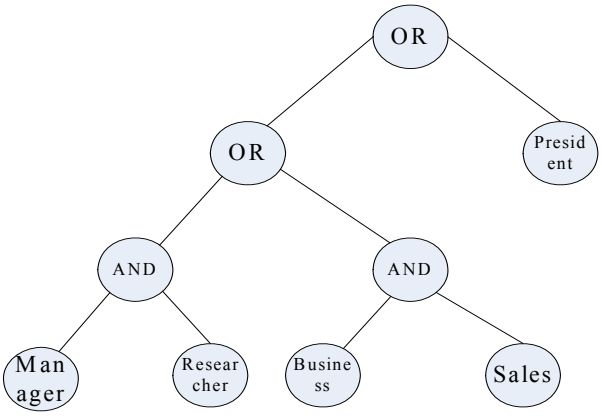
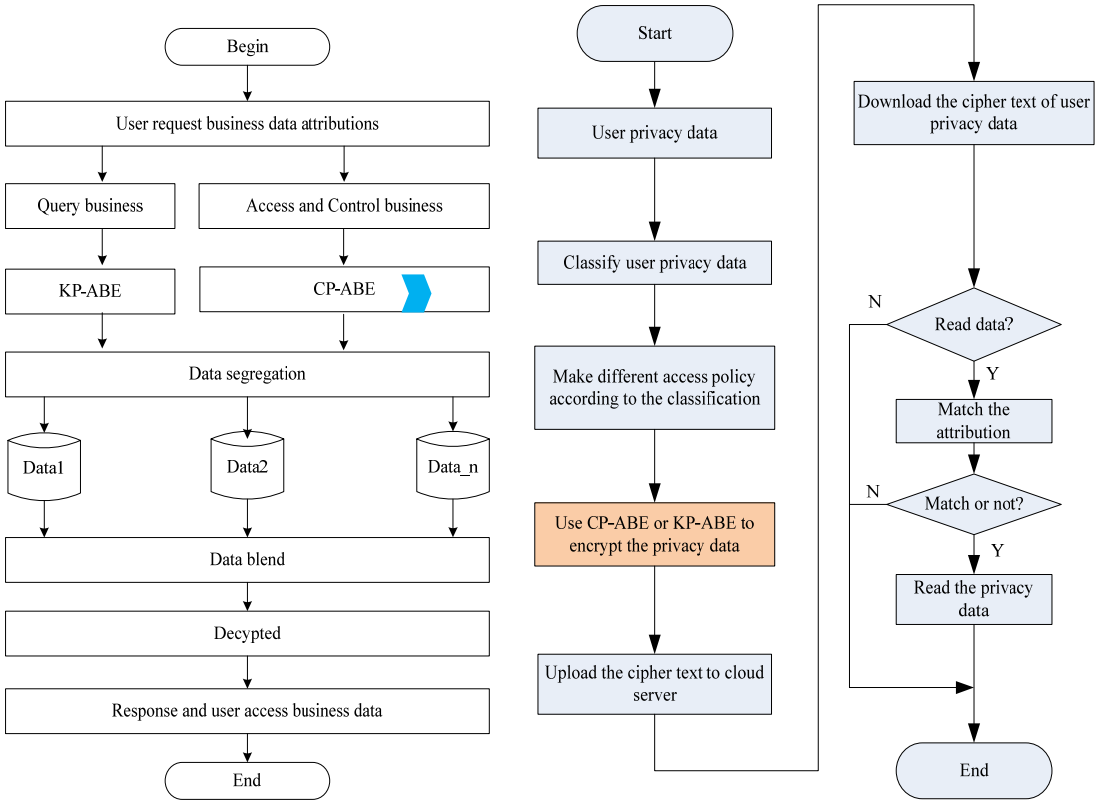


Figure 5. Low level ACT structure.

IV. The Case of Privacy Protection and its Implementation

A. Hybrid Solution for the Balance between Privacy Protection and Platform Performance in Cloud Computing

The hybrid solution proposed in this study differs from that in Ref. [13]. The dual solution in Ref. [13] uses both KP-ABE with subjective user attributes and CP-ABE with objective user attributes. The proposed solution chooses either CP-ABE or KP-ABE depending on customized user requirements. The choice of solution must promote a balance between performance and security. If the goal of the customer is to rapidly increase business, KP-ABE is chosen. Otherwise, CP-ABE is used. In cloud computing, access control businesses require high security and thus prefer CP-ABE; by contrast, query businesses require high performance and thus prefer KP-ABE [14]. This setup is referred to as the hybrid solution process, which is shown in Fig. 6(1). The implementation of the encryption and decryption process is illustrated in Fig. 6(2).



(1) Hybrid solution (2) Encryption and Decryption process

Figure 6. Hybrid solution for balancing cloud performance and security.

B. Privacy Protection in Cloud Computing

To verify the solution, a private cloud environment and platform is built and deployed in a selected campus. In this case, the private cloud environment plays five roles (Fig. 7): cloud service provider, user; universal description, discovery, and integration; key published center, and key distributed center. After uploading the data, the user obtains the key from the key distributed center. The cloud service provider receives the user data and encrypts the privacy information with ACP. Cipher text C can be read or written when the user can decrypt the privacy data with the private key.

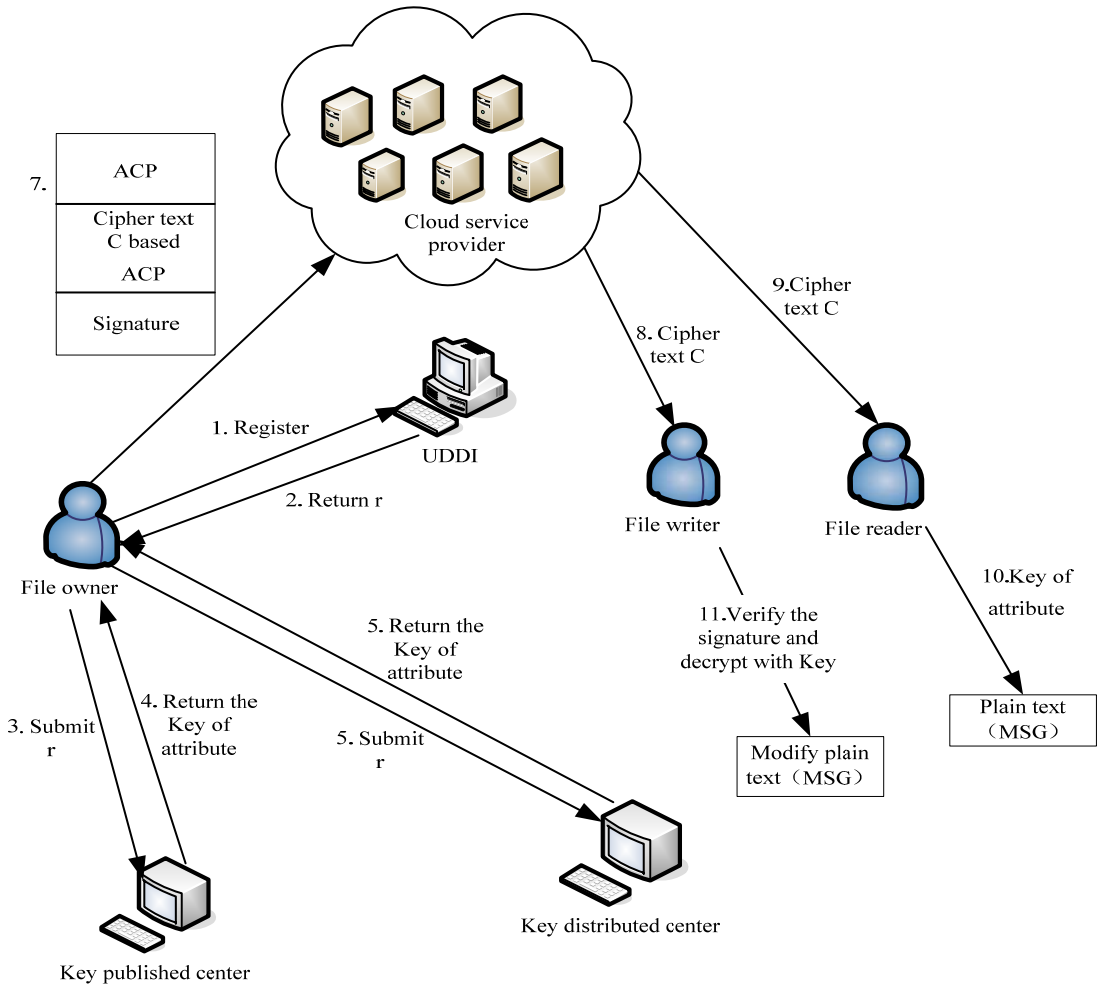


Figure 7. Privacy protection structure in cloud computing.

In this study, a customer relationship management (CRM) cloud service for telecoms is developed and deployed in the cloud environment (Fig. 7). The user-submitted personnel information is encapsulated into messages, which are then sent to an attribute-based key distribution center. The cloud environment then feeds back the public key and ACP to the user. Subsequently, the cloud computing and cloud storage platforms utilize both CP-ABE and KP-ABE to encrypt user privacy data. The detailed interaction between users and the cloud platform is shown in Fig. 8.

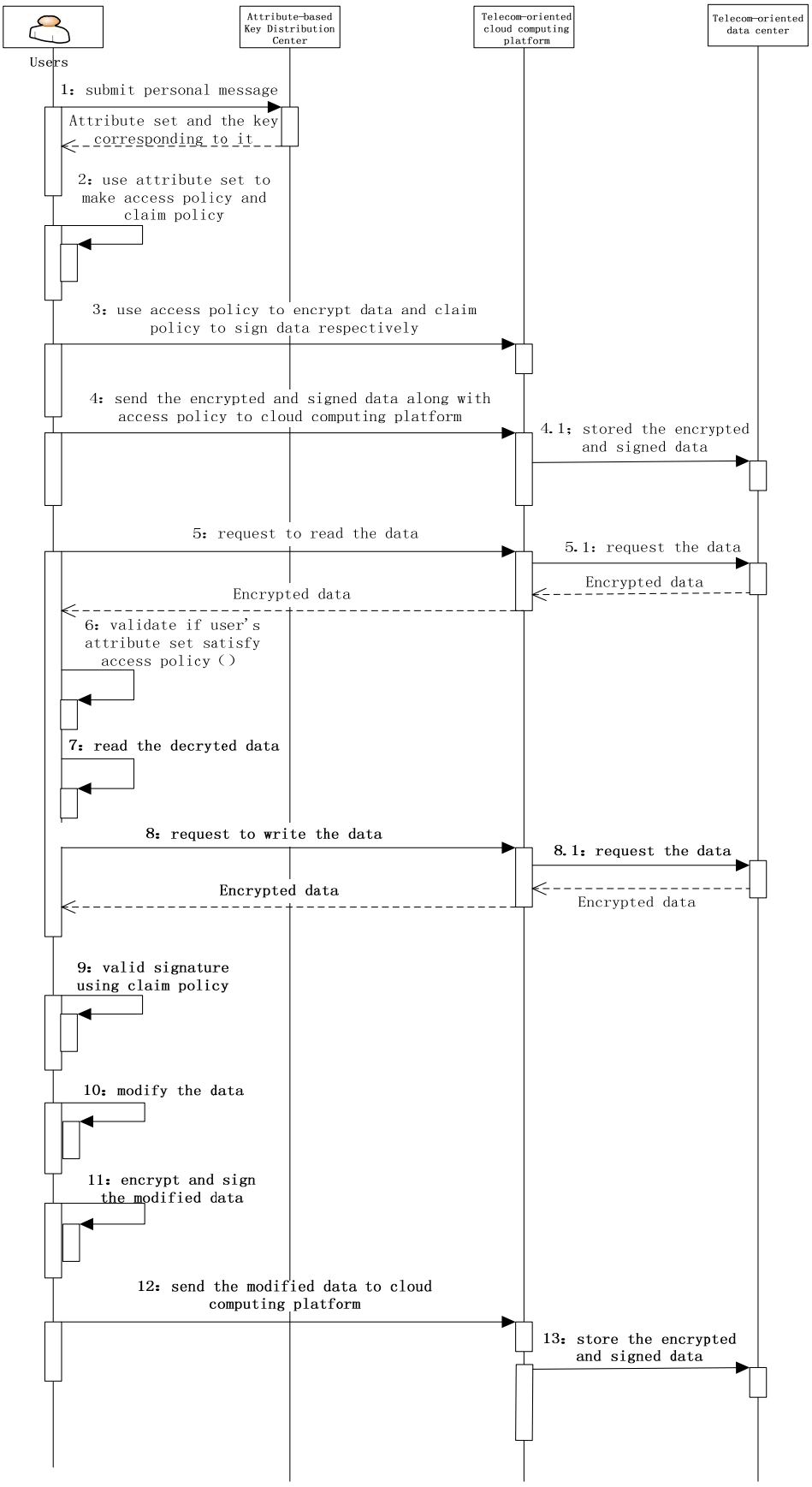


Figure 8. Interaction between users and the cloud platform.

C. Performance Analyses of Privacy Protection Solution in Cloud Computing

A CP-ABE toolkit was proposed in Ref. [15], and a KP-ABE toolkit was presented in Ref. [16]. The innovation of the hybrid solution is its capability of choosing between KP-ABE and CP-ABE during encryption of privacy data with ACT comprising user attributes. To assess the performance of the hybrid solution, a CRM cloud service called customer loss analysis is designed and developed as a product in the cloud computing portal shown in Fig. 9.

Figure 9. Portal of cloud service platform.

The customer information can be accessed by those with rights according to their attributes. In this case, the performance and security of the platform are both considered and implemented by the hybrid solution. This cloud service performance and security index is further improved and balanced, and the performance index is counted with an open cloud platform named Hadoop, as shown in Fig. 10.

Figure 10. The resources monitored of cloud environment based on Hadoop.

A CP-ABE toolkit is used to carry out the performance and security analysis shown in Fig. 10. The correlation of the cpabe-keygen, cpabe-en, and cpabe-dec attributes in the CP-ABE toolkit with the number of attributes is represented by a straight line when the toolkit is deployed and run.

V. Conclusion

Although the use of the proposed solution in balancing performance and security in cloud computing shows great promise, a large amount of data are needed to verify its effectiveness. The next work will focus on big data in cloud computing and will continue to verify and optimize the hybrid solution.

References

[1] Demchenko Y., A. Mavrin & C. de Laat (2011). Defining generic architecture for cloud infrastructure as a service provisioning model. CLOSER2011 Conference. Nordwijk, Netherlands.

[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Kon-winski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica & M. Zaharia (2009). Above the clouds: A Berkeley view of cloud computing. University of California, Berkeley. Tech. Rep. UCB-EECS-2009-28.

[3] Gellman, R. (2009). Privacy in the clouds: risks to privacy and confidentiality from cloud computing. World Privacy Forum.

[4] Mowbray, M. & S. Pearson (2009). A client-based privacy manager for cloud computing. COMSWARE'09.

[5] H. Li, Y. Dai, L. Tian & H. Yang (2009). Identity-based authentication for cloud computing. In CloudCom, ser. Lecture Notes in Computer Science (vol. 5931). Springer, 157–166.

[6] C. Gentry (2009). A fully homomorphic encryption scheme. Ph.D. dissertation. Stanford University. http://www.crypto.stanford. edu/craig.

[7] Sadeghi, A.R., T. Schneider & M. Winandy (2010). Token-based cloud computing. In TRUST, ser. Lecture Notes in Computer Science (vol. 6101). Springer, 417–429.

[8] G. Wroblewski (2002). General method of program code obfuscation. Ph.D. dissertation. Wroclaw University of

Technology. http://www.ouah. org/wobfuscation.pdf.
[9] Sahai, A. & B. Waters (2005). Fuzzy identity-based encryption. In *Advances in Cryptology*. Eurocrypt (vol. 3494) of LNCS, 457–473. Springer.
[10] Fang, W., B. Yang & D. Song (2010). Preserving private knowledge in decision tree learning. Journal of Computers (vol. 5, no. 5), 733–740.
[11] Bethencourt, J., A. Sahai & B. Waters (2007). Ciphertext-policy attribute-based encryption. In IEEE Symposium on Security and Privacy, 321–334.
[12] Goyal, V., O. Pandey, A. Sahai & B. Waters (2006). Attribute-based encryption for fine-grained access conrol of encrypted data. In ACM Conference on Computer and Communications Security (ACM CCS).
[13] Attrapadung, N., H. Imai (2009). Dual-policy attribute based encryption [C]. Proceedings of Applied Cryptography and Network Security. Lecture Notes in Computer Science (vol. 5536). Berlin: Springer-Verlag, 168–185.
[14] Yubin, G., Z. Liankuan, L. Fengren & L. Ximing (2013). A solution for privacy-preserving data manipulation and query on nosql database. Journal of Computers (vol. 8, no. 6), 1427–1432.
[15] http://hms.isi.jhu.edu/acsc/cpabe/.
[16] http://sourceforge.net/p/kpabe/code/1/tree/com/zaranux/crypto/abe/kp/.

Yi-mu Ji was born in Anhui Province, China on September, 978. He received doctoral degree in Information and Telecommunication System at Nanjing University of Posts and Telecommunications in 2007. Now he is an association professor and post-doctoral in Nanjing University of Posts and Telecommunications, and he is interested in the researching fields of Computer software theory, network and cloud computing.

Jie Tan was born in Jiangsu Province, China on September,1990. Now he is a post-graduation student in Nanjing University of Posts and Telecommunications, and he is interested in cloud computing and SDN routing protocol.