TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 09/10 pp738–749
DOI: 10.26599/TST.2018.9010127
Volume 24, Number 6, December 2019

Cloud Storage Security Assessment Through Equilibrium Analysis

Yuzhao Wu, Yongqiang Lyu, and Yuanchun Shi*

Abstract: With ever greater amounts of data stored in cloud servers, data security and privacy issues have become increasingly important. Public cloud storage providers are semi-trustworthy because they may not have adequate security mechanisms to protect user data from being stolen or misused. Therefore, it is crucial for cloud users to evaluate the security of cloud storage providers. However, existing security assessment methods mainly focus on external security risks without considering the trustworthiness of cloud providers. In addition, the widely used third-party mediators are assumed to be trusted and we are not aware of any work that considers the security of these mediators. This study fills these gaps by assessing the security of public cloud storage providers and third-party mediators through equilibrium analysis. More specifically, we conduct evaluations on a series of game models between public cloud storage providers and users to thoroughly analyze the security of different service scenarios. Using our proposed security assessment, users can determine the risk of whether their privacy data is likely to be hacked by the cloud service providers; the cloud service providers can also decide on strategies to make their services more trustworthy. An experimental study of 32 users verified our method and indicated its potential for real service improvement.

Key words: cloud storage security; security assessment; equilibrium analysis

1 Introduction

Cloud storage services have been widely used and an increasing number of end users, organizations, and enterprises are storing their personal or business data in the cloud. In perhaps the simplest cloud storage model, data are uploaded to the cloud, and data consumers (possibly different from the owner of the data) access the data through cloud servers. Such a model is very convenient to realize large-scale, multi-regional, and multi-domain data sharing. However, as users' data may contain sensitive information related to privacy or secrets, the third-party cloud storage Service Providers (SPs) holding large amount of data are becoming the targets of network attackers and some SPs are becoming the data abusers themselves. These security and trust issues are of great importance in the big data era when data carries great values. It is equally important to develop security assessment methods to support users to make good choices on public cloud storage service from the security stand point and to help service providers to evaluate the security of their services.

A number of frameworks and schemes have been proposed for the risk assessment of cloud computing to evaluate the security of cloud platforms. For example, the authors in Refs. [1, 2] developed their own cloud security assessment frameworks based on risk management assessment methods adopted from Ref. [3]. Those studies usually treated cloud providers as defenders of security and assumed them to be

• Yuzhao Wu are with the Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China. E-mail: wuyz11@mails.tsinghua.edu.cn.
• Yongqiang Lyu are with the Research Institute of Information Technology & TNList, Tsinghua University, Beijing 100084, China. E-mail: lvyq@tsinghua.edu.cn.
• Yuanchun Shi are with the State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China. E-mail: shiyc@mail.tsinghua.edu.cn.
∗ To whom correspondence should be addressed.
Manuscript received: 2018-07-05; accepted: 2018-09-25

trustworthy. However, public cloud providers are only semi-trustworthy, which means they are honest but curious: they will honestly execute the operation we call them for, but at the same time may acquire the content of data. In alternative frameworks, the roles of the cloud providers are distinguished. According to Ref. [4], there are two primarily types of cloud providers: Cloud Service Providers (CSP) or SaaS or PaaS providers, e.g., Google App Engine, which offer cloud services over the Internet; and Cloud Infrastructure Providers (CIP) or IaaS providers, e.g., Amazon EC2, which provide cloud infrastructures (typically virtualized execution environments) as a service and thus serve as the foundation layer for cloud systems. The two types of cloud providers have different security roles and may face conflicts. According to Ref. [5], a CSP is responsible for nearly all of the security requirements because in this model both data access and computation are done on the provider side, whereas a CIP is generally responsible only for the availability of resources but not for security. To the contrary, Wazir et al.[6] pointed out that CSPs face difficulty in providing security for user data to ensure confidentiality, integrity, reliability, availability, and privacy. These studies showed that there usually exist conflicts of benefits between attackers and defenders, different layers of cloud providers, as well as cloud providers and users. These conflicts of benefits affect the behavior decisions of cloud providers and possibly render them semi-trustworthy to users, who still lack adequate assessment mechanisms.

In order to address the semi-trustworthiness of cloud providers, third-party mediators are often installed on the access control management frameworks of cloud storage[7–9]. They may be verifiers for the cloud providers, or encryption servers managing cloud users' private data. There are some Third-party Service Providers (TSPs) acting as mediators to offer security services in current commercial cloud-based applications. Although these TSPs are assumed trustworthy, they may also face benefit conflicts with cloud providers and users, which makes them also semi-trustworthy, similar to the cloud providers. Security assessment considering the benefit conflicts related to TSPs is also lacking.

The most widely-used method to solve benefit conflicts is game theory. Game theory could be used naturally as a defensive measure because during independent and strategic rational decision making, each cloud user will compete for their own best possible solution[10]. Many security evaluation and risk assessment schemes have been proposed based on game theory. Manshaei et al.[11] summarized the game theory approaches towards different topics of security, mainly including security at the physical and MAC layers (e.g., jamming and eavesdropping attacks), security of self-organizing networks (e.g., revocation in mobile ad hoc networks), intrusion detection systems (e.g., collaborative IDS), anonymity and privacy (e.g., cooperative location privacy), economics of network security (e.g., interdependent security), and cryptography (e.g., security in multi-party computation).

From the viewpoint of benefits, the problem of cloud storage security is not only a technical problem but also an economic one. The economics of network security is therefore another active research topic. Researchers have already investigated dependability and software economics, behavioral economics, and the psychology of security for analyzing and solving certain security and privacy problems[12–14]. Game theory has also been one of the main tools used to analyze the economics of security.

In the research into interdependent security, security can be viewed as a good; everyone benefits when the network guarantees security and everyone suffers otherwise[11]. The security of the whole system depends on the collective behavior of nodes in the network. According to Refs. [15, 16], in interdependent systems, each individual's benefit is determined by the average security level of the whole system. In these models, an attacker needs to conquer a majority of the machines in the network one-by-one to succeed in its attack goal.

The problems in security fields usually involve decisions among multiple layers, thus game theoretic models can be constructed on them. The Decision Makers (DMs) can often be divided into attackers and defenders, who have contrary benefits. Attackers attempt to break and interrupt systems and defenders try to prevent and protect the systems from suffering damages. Moreover, in cloud computing environment, there are various types of DMs including users and different layers of cloud providers. Users and cloud providers can be defenders as well as attackers, depending on the benefits. Game theory can offer mathematic tools and models for DMs to decide their strategies. One widely used game theoretic model is the Nash equilibrium. If a strategy profile of a game makes each DM cannot individually benefit more from a change of strategy, this strategy profile is referred to one of the game's Nash equilibrium. Along with its extended forms, it is an effective method to model security games between attackers and defenders. For example, Ref. [17] uses the Nash equilibrium to analyze a Stackelberg competition between one attacker and multiple defenders, and Ref. [18] uses a generalized Nash equilibrium to analyze the game between a CSP and a CIP, pointing out that a CSP can make a CIP share the security risk by employing certain strategies. However, these studies do not provide a clear assessment on cloud security because of a lack of metrics and evaluation theory.

In this study, we focus on assessing the security of public cloud storage providers and third-party mediators through equilibrium analysis. This study can help both cloud users and service providers to make better choices and to benefit more from security investment. The main contributions of our work are as follows.

— We build multiple game-theoretic models between cloud users and providers and implement an equilibrium analysis method on the security games. Our study covers one-user, multi-user, and multi-service provider models.

— We analyze several known cloud storage frameworks using our proposed models to evaluate the advantages and disadvantages for their security. More specifically, we have developed a series of theorems and security guidance for users, CSPs, and TSPs to make better choices regarding security.

— We conduct a real user study to verify our theory and methods. The results suggest that our work has great potential for the security improvement of real systems.

2 Related Work

2.1 Work on access control management

Many researchers have examined access control management as a means to improve the security of traditional database environments. Access control is applied over the most important problems. Cloud storage systems, which can be thought of as databases in the cloud, have also attracted much research attention to access control management.

Bertino and Ferrari[19], as well as Gerome and Dan[20], tried a key pre-distribution scheme in untrustworthy storage. This scheme does not have good scalability and thus cannot support fine-grained access control. Di Vimercati et al.[21] gave to data a fine-grained encryption based on an Access Control List (ACL), in which a user can only save one key to derive all the authorities he needs; but this will become more expensive as the number of users increases. Goyal et al.[22] first constructed a Key-Policy Attribute-Based Encryption (KP-ABE), applying the idea of identity-based encryption proposed by Shamir[23] to access control in the cloud. Wang et al.[24] put forward a hierarchical management of key policy based on Ciphertext-Policy Attribute-Based Encryption (CP-ABE), but the hierarchical management only applied to storing information on users and attributes. Nabeel et al.[25–27] constructed, then improved upon a group key management scheme on broadcast called Access Control Vector Broadcast Group Key Management (ACV-BGKM), which uses an access tree to give the scheme greater advantages. They resolved the policies to both the owner and the cloud to make it a two-layer key policy, but a single entity's power was weakened and it was prone to collision attacks.

The works above make sure that servers can only access data which have been encrypted by users, thus ensuring the inner security of the system. Nevertheless, the access control management increases the cost of data exchange, and cannot prevent cloud providers from decrypting user data. More recent studies have preferred to perform more encryption and decryption operations in the cloud and use third-party providers to deal with security. For example, Li et al.[7] offloaded most of the key generation related operations to a key update CSP. Yi et al.[8] considered a scenario where a cloud user outsourced the task to encrypt and store its data to several "semi-honest" SPs. These SPs are in relations of cooperation, as we mentioned in Section 2.3 above. Additionally, Yong et al.[9] suggested a verifier to check if the SP is actually storing data honestly and Sharma and Joshi[28] used a third-party "reasone" to decide what is permitted in an ABE model using Web Ontology Language (OWL).

These studies have used different kinds of third-party mediators for their cloud storage access control management models, showing that the integrity of CSPs is in doubt and that they believe TSPs are more secure than CSPs, even though there is a lack of proof for this belief.

2.2 Work on cloud security assessment

Cuschieri[2] proposed a cloud security risk assessment framework including 11 processes, which is similar to the work of the Cloud Security Alliance[1] and based on ISO 31000[3]. The framework can be summarized as Fig. 1.

Fitó et al.[4] proposed a risk assessment approach based on Business-Level Objectives (BLO), which enables a CSP to maximize its profit by transferring the risks of provisioning its private cloud to third-party CIPs. Sangroya et al.[29] analyzed the advantages and disadvantages of cloud computing and surveyed the security mechanisms of the major cloud platforms, concluding with a risk assessment scheme to be employed by CSPs. Kalishi Jr and Pauley[30] and Theharidou et al.[31] proposed the concept of Assessment-as-a-Service (AaaS), which suggested making risk assessment a necessary service of cloud platforms, but they did not implement the proposal.

From the quantity of work, we can see that the cloud security problem has attracted enough attention. However, in the cloud security assessment field, most of the current work is for helping cloud users assess their risk before putting their critical data in a security sensitive cloud. All of this research has laid a solid foundation for cloud computing, but it has not added up to a complete risk assessment approach in consideration of the specific and complex characteristics of cloud computing environments[32]. These frameworks can make a reasonable assessment on the risk of the cloud platform being under attack, but they do not consider the risk that cloud providers abuse users' data.

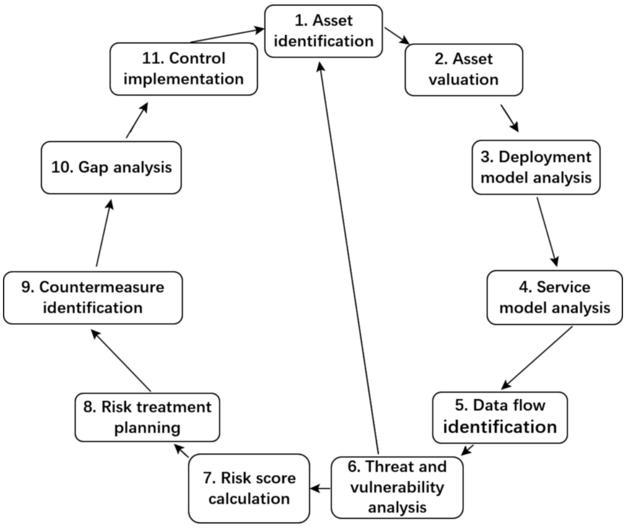
1. Asset identification
2. Asset valuation
3. Deployment model analysis
4. Service model analysis
5. Data flow identification
6. Threat and vulnerability analysis
7. Risk score calculation
8. Risk treatment planning
9. Countermeasure identification
10. Gap analysis
11. Control implementation

Fig. 1 A cloud security risk assessment framework from Ref. [1].

2.3 Work on implementing game theory in cloud security

Game theory has been widely used in the security field. As introduced above, Lou and Vorobeychik[17] and Ardagna et al.[18] analyzed the security games between attackers and defenders and between CSPs and CIPs, respectively, to find the Nash equilibrium or generalized Nash equilibrium. Further, Furuncu and Sogukpinar[5] not only analyzed the game and equilibrium between multiple attackers and defenders, but also established a scalable Risk Assessment Method for Cloud Computing using game theory (CCRAM). They used the standards in the National Vulnerability Database (NVD), in which the impacts of attackers are divided into Confidentiality, Integrity, and Availability (CIA). Based on the goal of the attackers, they divided attacks into categories such as DoS (Denial of Service), user (gaining user privilege), data (non-permitted data access and write), administrative (gaining administrator user privilege), and scan (getting information about the target system). CCRAM is an imperfect information non-cooperative non-zero static game model, which makes a lot of assumptions. It assumes that among multiple attackers and defenders, the attacks will succeed as long as defenders do not defend and fail as long as some defenders defend. It also assumes that there is no cooperation between attackers and defenders, and their behavior will all bring costs such that it is a non-zero game. Meanwhile they implement a number of parameters based on these assumptions to get a mixed Nash equilibrium. Their conclusions cannot be universalized, but their research methods are reasonable.

Ismail et al.[33] performed a game theoretical analysis on the interaction between the verifier and the cloud provider, in which they formulated the problem as a two-player non-cooperative game. Reference [10] presents a review of many possible security threats and their countermeasures using game-theoretic approaches.

The majority of works on implementing game theory in the cloud security field focus on games between attackers and defenders, or between CSPs and CIPs, but few focus on the security game between cloud users and providers. Moreover, the security of TSPs has never been analyzed.

3 Modeling Cloud Security Games

In order to describe the game between cloud users and providers, we proceed with our modeling effort in three steps. The three models reveal an increasing number of properties and approaches as the complexity of the situations grows.

3.1 One-user model

At first, we model the game between one user and one CSP. In this model, the user makes a decision between using the cloud service or not, and the CSP makes decision between stealing the user's private data or remaining honest. The utility functions of both sides only depend on the two DMs and are not affected by other users on the cloud.

We use Table 1 to present a standard form of the game between a user i and a CSP c.

If the user i chooses to use the cloud, he will gain some benefits while paying some fees, but if his data was to be stolen by the CSP, he would suffer damages. If the CSP c chooses to steal the user's data, he could gain some additional benefits above the fees paid by the user choosing to use the cloud. But if the CSP's theft were discovered, he would suffer damages through punishment.

Here we state some assumptions behind the game:

(1) If the user chooses to use the cloud, he could receive profit C equal to the data's value, while he would lose the same amount of profit C if his data were to be stolen. If he then discovered that the CSP was stealing, he could expose the provider and receive some compensation Q, where $Q \leqslant C$.

(2) If the CSP chooses to steal the user's data, he could receive profit C equal to the data's value and thus equal to the profits the user could obtain from using the cloud. If his stealing behavior was discovered and exposed, he would be punished and pay P, where $Q \leqslant P$. That is to say, he may need to pay a fine on top of the compensation he would pay to the user.

(3) If the user chooses to use the cloud despite not trusting the CSP, he could encrypt the data before putting it in the cloud. The encryption would bring an encryption cost E to the user and a decryption cost to the CSP who would like to steal the data. The decryption cost is positively correlated to the encryption cost, which can be expressed as $F(E)$.

Set the user i's data value to be C_i, the encryption cost to be E_i, the fees paid to the CSP to be M_i, the probability that the user finds the CSP stealing to be p_f and the compensation the user could receive if the CSP is discovered to be stealing to be Q_i. When the user i chooses to use the CSP c and the CSP chooses to steal the data, the utility function of the user i is

$$B_i = C_i - E_i - M_i - C_i + p_f Q_i = p_f Q_i - (E_i + M_i) \tag{1}$$

When the user i chooses to use the CSP c and the CSP chooses not to steal the data, the utility function of the user i is

$$B_i' = C_i - E_i - M_i = C_i - (E_i + M_i) \tag{2}$$

Set the CSP c's decryption cost for user i's data to be $F(E_i)$, the exposure probability of stealing to be p_f and punishment for exposure to be P_i. When the user i chooses to use the CSP c and the CSP chooses to steal the data, the utility function of the CSP c is

$$B_c = M_i - F(E_i) - p_f P_i + C_i \tag{3}$$

When the user i chooses to use the CSP c and the CSP chooses not to steal the data, the utility function of the CSP c is

$$B_c' = M_i \tag{4}$$

When the user i chooses not to use the CSP c, the utility of both the user i and CSP c is 0.

Table 1 Standard form of the game between user and cloud.

	Do not steal users' data	Steal user's data
Use the cloud	(B_i', B_c')	(B_i, B_c)
Do not use the cloud	(0, 0)	(0, 0)

3.2 Multi-user model

Further, we consider a game model with multiple users and one CSP. In such a model, the CSP needs to make decisions on whether to steal the data for each user that chooses his service and in consideration of his total profits. For users, although their benefits from using the cloud are independent, if one user was to find that the CSP stealing his data, all the users would lose their trust in the CSP and check whether their own data had been stolen.

Thus in this model, as long as one user finds the CSP stealing his data, all of the users whose data was stolen would discover the theft and call for compensation. In order to analyze the exposure probability, we set a trust degree ρ which is among [0, 1] to represent the degree of trust that the CSP has with users. Then the exposure probability p_f of the CSP should satisfy the conditions below:

(1) p_f increases as the trust degree ρ decreases; when $\rho = 0$, $p_f = 1$; when $\rho = 1$, $p_f = 0$.

(2) p_f increases when the value of the data C that the CSP has stolen increases. When $C = 0$, the stolen

data has no value, and we can consider the CSP faces no risks or costs, thus $p_f = 0$. When $C \to \infty$, $p_f = 1$.

(3) When ρ is fixed, the exposure probability when the CSP steals multiple users' data is equal to the exposure probability when the CSP steals a single users' data the value of which is the same as the total value of the multiple users' data.

To satisfy the three conditions above, we can express the exposure probability p_f for a CSP with trust degree ρ to steal data whose value is C by the formula below:

$$p_f = 1 - \rho^C \tag{5}$$

We can see that it satisfies conditions (1) and (2) easily.

For condition (3), we consider two cloud users i and j, whose data values are C_i and C_j, respectively. We can obtain the probability that they discover that the CSP is stealing from $p_{f_i} = 1 - \rho^{C_i}$ and $p_{f_j} = 1 - \rho^{C_j}$, respectively. Thus we can compute the probability that the CSP is exposed by the two users as

$$\begin{aligned} p_{f_{i,j}} =& 1 - \left(1 - p_{f_i}\right)\left(1 - p_{f_j}\right) = \\ & 1 - [1 - \left(1 - \rho^{C_i}\right)][\left(1 - \rho^{C_j}\right)] = \\ & 1 - \rho^{C_i}\rho^{C_j} = 1 - \rho^{C_i + C_j} \end{aligned} \tag{6}$$

This probability is equal to the exposure probability of the CSP with a user whose data value is $C_i + C_j$, thus we can see the formula satisfies condition (3).

After introducing the trust degree and exposure probability, we come back to the utility function of the multi-user model. As with the one-user model, we can get the utility function of the user i when he chooses to use the CSP c and the CSP chooses to steal the data as

$$B_i = p_f Q_i - (E_i + M_i) \tag{7}$$

Here p_f is the CSP's exposure probability for all users, not only the user i.

When the user i chooses to use the CSP c and the CSP chooses not to steal the data, the utility function of the user i is

$$B_i' = C_i - (E_i + M_i) \tag{8}$$

The CSP c would make a decision for each user as to whether to steal their data or not. We use the steal probability p_i to express whether c would steal user i's data. When the CSP chooses to use a pure strategy, meaning he has just two choices, stealing with the probability 100% or 0%, we can set p_i as 1 and 0.

Thus the utility function of the CSP c on all the users is

$$B_c = \sum M_i - \sum p_i F(E_i) - p_f \sum P_i + \sum p_i C_i \tag{9}$$

Especially, when the CSP chooses not to steal all the users' data, $\forall\ p_i = 0$, the utility is

$$B_c' = \sum M_i \tag{10}$$

3.3 Multi-SP model

Finally, we consider models in which cloud users can choose among different service providers.

In these models, a user's utility function depends on the service providers he chooses and the other users who shared the same cloud, thus it is not changed from the utility function in the multi-user model.

But when taking account of multiple SPs, we need to consider that they have different kind of relationships. Basically, we can divide their relationships into three types: competitive, cooperative, and dependent. In competitive relations, different SPs maintain their services and conditions individually. In cooperative relations, the user employs the resources of multiple SP in a distributed manner. In dependent relations, some of the SPs serve only as third-party SPs providing security services. These TSPs are dependent on other CSPs who hold the cloud servers. In this relation, users will choose to either use a CSP's service directly or to make use of a TSP to maximize data security.

There are different types of game theory models for different kinds of relationships. In competitive and cooperative relations, games are usually among different SPs. If an SP is disconnected by other SPs or not chosen by the user, it will not benefit. In dependent relations, on the other hand, the CSP can benefit regardless of whether the user chooses the CSP itself or a TSP who is dependent on the CSP. In this situation, the game models are more focused on security. We will make a further analysis of this relation in Section 5.

4 Case Study with One-User Model

Our first result lies in the one-user model and its Nash equilibrium.

As we have shown above, when user i chooses to use the cloud and CSP c chooses to steal data, the utility functions of the i and c are Eqs. (1) and (3); and when user i chooses to use the cloud and CSP c chooses not to steal, the utility functions of the i and c are Eqs. (2) and (4).

When user i chooses not to use the cloud, the utilities of i and c are both 0.

According to the assumptions in the model, $Q_i \leqslant C_i$, thus we have

$$B_i = p_f Q_i - (E_i + M_i) \leqslant C_i - (E_i + M_i) = B_i'.$$

That is to say, when the CSP chooses to steal the data, the profits of the user will decrease.

Theorem 1 In the one-user model, the user must encrypt his data in a circumstance where there is no punishment for the CSP when it steals the data, and the CSP would not steal the user's data if the decryption cost is greater than the data's value.

Proof Because M is the fees that the user paid to the CSP, it cannot be more than the user's data value C, otherwise $B'_i = C_i - (E_i + M_i) \leqslant C_i - M_i \leqslant 0$, the user would not choose the cloud.

When $M \leqslant C$, we have

$$B_c - B'_c = M_i - F(E_i) - p_f P_i + C_i - M_i = C_i - p_f P_i - F(E_i) \tag{11}$$

If the user does not encrypt his data, $F(E_i) = 0$ meanwhile $P_i \leqslant C_i$, then $B_c - B'_c$ is not less than 0. That is to say, when the CSP chooses to steal, he can ensure his profits are more than from not stealing. If we do not have an effective punishment mechanism, the CSP is untrusted. In this situation, the user must encrypt his data to ensure security.

When $C_i \leqslant F(E_i)$, $B_c - B'_c \leqslant 0$, the CSP would not benefit from changing his strategy to steal. Thus Theorem 1 is proved. ■

5 Case Study with Multi-User Model

Here we analyze the model between multiple users and one CSP.

In the multi-user case, when user i chooses to use the cloud and CSP c chooses to steal, the utility function of i is Eq. (7); when user i chose to use the cloud and CSP c chose not to steal, the utility function of i is Eq. (8). The utility function of the CSP c is Eq. (9). Especially, when the CSP chose not to steal all the users' data, $\forall\ p_i = 0$, the utility is Eq. (10).

When the CSP chooses to steal user i's data, his exposure probability on i is $p_f = 1 - \rho^{C_i}$. From the function's properties, we can calculate the total exposure probability of CSP as

$$p_f = 1 - \rho^{\sum p_i C_i} \tag{12}$$

Theorem 2 In the multi-user model, even though some users do not encrypt their data, there remains a possibility that the CSP would choose not to steal their data.

Proof Assume that the CSP's punishment for stealing is equal to the total fees that all users have paid to him, which is to say, $P_i = M_i$. Then

$$\begin{aligned} B_c = & \sum M_i - \sum p_i F(E_i) - p_f \sum M_i + \sum p_i C_i = \\ & (1 - p_f) \sum M_i + \sum p_i [C_i - F(E_i)] = \\ & \rho^{\sum p_i C_i} \sum M_i + \sum p_i [C_i - F(E_i)] \end{aligned} \tag{13}$$

Thus

$$\begin{aligned} & B_c - B'_c = \\ & \rho^{\sum p_i C_i} \sum M_i + \sum p_i [C_i - F(E_i)] - \sum M_i = \\ & \sum p_i [C_i - F(E_i)] - (1 - \rho^{\sum p_i C_i}) \sum M_i \end{aligned} \tag{14}$$

Because $\rho \leqslant 1$, we can get that if $C_i \leqslant F(E_i)$,

$$B_c - B'_c \leqslant \sum p_i [C_i - F(E_i)] \leqslant 0.$$

That is to say, when the decryption cost for some users is higher than the data value, the CSP cannot obtain a benefit from stealing these users' data. Therefore, the best strategy for the CSP is not to steal these particular users' data.

When $C_i > F(E_i)$, we assume the CSP would not steal any users' data by default, if CSP changed his steal probability on one single user j to p_j, we have

$$B_c^j = \rho^{p_j C_j} \sum M_i + p_j [C_j - F(E_j)] \tag{15}$$

Comparing to the utility B'_c when the CSP does not steal anyone's data,

$$B'_c - B_c^j = (1 - \rho^{p_j C_j}) \sum M_i - p_j [C_j - F(E_j)] \tag{16}$$

There exists a minimal K such that when $K \geqslant C_i - F(E_i)$, i.e., $F(E_i) \geqslant C_i - K$, we have $B'_c - B_c^j \geqslant 0$. In this situation, the CSP would not steal users' data.

That is to say, if user i made his data's decryption cost to be higher than $C_i - K$, the CSP would not steal his data. Also we should notice that if user i did not encrypt his data, where $F(E_i) = 0$, then $B'_c - B_c^j = (1 - \rho^{p_j C_j}) \sum M_i - p_j C_j$. We can regard it as a function $f(p_j)$ on p_j, then $f(p_j) = 0$ is a transcendental equation with a root $p_j = 0$. According to the value of $\sum M_i$ and C_j, $f(p_j)$ can be positive in the range (0, 1), also it can be lower than 0 in the same interval. Therefore, in the multi-user model, even if some users did not encrypt their data, if there exists a Nash equilibrium choice for the CSP in which he would not steal any users' data, the unencrypted data's security can also be ensured. This security comes from the encryption work of other users. ■

6 Case Study with Multi-SP Model

Here we analyze a multi-SP model constructed from the general TSP security platform framework.

A TSP does not own the cloud resources, merely providing a security service for users who do not trust the CSP and need more access control abilities for their data. The relation of TSP to CSP is similar to the relation of CSP to CIP, where the service of the former is based on the latter. In the multi-SP model, users choose between TSP and CSP, where the utility of TSP is influenced by CSP. Whatever the users choose, the CSP's cloud resources are acquired. Therefore, the utility that CSP can obtain from providing his cloud storage service is unchanged.

Figure 2 shows a typical TSP framework now in use. When the user chooses to use the cloud service directly, he would encrypt the private data himself and upload it to the cloud, then decrypt the data each time it was downloaded. When he chooses to use the service provided by a TSP, he can outsource all of the encryption and decryption work to the third-party platform as long as it is trustworthy. In this situation, the user can gain in terms of usability and convenience while not sacrificing security.

As a matter of fact, however, the TSP is semi-trustworthy just as the CSP is. Users who choose the TSP should encrypt their data first thus gaining nothing more in terms of usability through the TSP, but gaining additional security.

Theorem 3 Choosing a TSP will provide a security gain over a CSP if the TSP and CSP have the same trust degree.

Proof Assume that the CSP requires payment M_i from user i while the TSP requires m_i and then puts all of the users' data on CSP's cloud platform, that's to say, TSP will pay M_i to CSP for user i's data. According to this assumption, we can get $m > M$, otherwise the TSP cannot benefit.

Set TSP's trustiness degree to be ρ_{TP}, we can get its utility function:

$$\begin{aligned}B_{\text{TP}} = &\sum(m_i - M_i) - \sum p_i F(E_i) - p_f \sum m_i + \\ &\sum p_i C_i = \\ &(1-p_f)\sum m_i - \sum p_i F(E_i) + \sum p_i C_i - \\ &\sum M_i = \\ &\rho_{\text{TP}}{}^{\sum p_i C_i} \sum m_i + \sum p_i [C_i - F(E_i)] - \\ &\sum M_i = \\ &\rho_{\text{TP}}{}^{\sum p_i C_i} \sum M_i + \sum p_i [C_i - F(E_i)] + \\ &\rho_{\text{TP}}{}^{\sum p_i C_i} \sum (m_i - M_i) - \sum M_i\end{aligned} \tag{17}$$

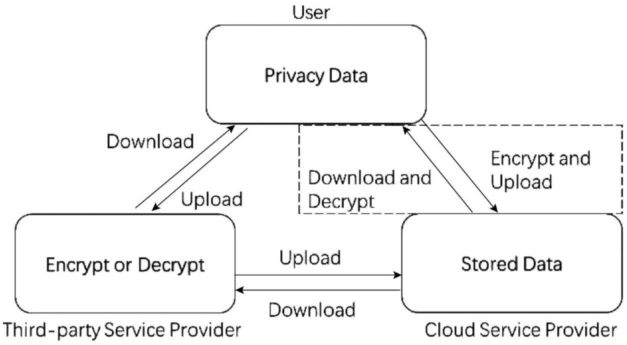

Fig. 2 Third-party security service platform framework.

Especially, when TSP chooses not to steal all the users 'data, $\forall\ p_i = 0$, we have

$$B'_{\text{TP}} = \sum(m_i - M_i) \tag{18}$$

Hence,

$$\begin{aligned}&B_{\text{TP}} - B'_{\text{TP}} = \\ &\rho_{\text{TP}}{}^{\sum C_i} \sum M_i + \sum p_i [C_i - F(E_i)] - \\ &(1 - \rho_{\text{TP}}{}^{\sum C_i}) \sum m_i - \rho_{\text{TP}}{}^{\sum C_i} \sum M_i\end{aligned} \tag{19}$$

When $\rho_{\text{TP}} = \rho$,

$$\begin{aligned}&(B_{\text{TP}} - B'_{\text{TP}}) = \\ &\rho^{\sum C_i} \sum M_i + \sum p_i [C_i - F(E_i)] - \\ &(1 - \rho^{\sum C_i}) \sum m_i - \rho^{\sum C_i} \sum M_i = \\ &(B_c - B'_c) - (1 - \rho^{\sum C_i}) \sum (m_i - M_i) < \\ &(B_c - B'_c)\end{aligned} \tag{20}$$

As a result, we can see that when there is an equal degree of trusts in the TSP and CSP, the TSP would obtain less profits than the CSP by changing strategy from not stealing the user data to stealing. Accordingly, there needs to be a higher value of K to make the TSP tend to steal users' data. In this situation, we can conclude that the TSP is more secure than the CSP for the users. ■

7 Experimental Results

In this section, we use an experimental method to test and verify our models.

7.1 Methods

We chose 32 cloud storage individual users as respondents to a questionnaire survey, of which 16 persons are employees and 16 persons are undergraduate or graduate students. On the premise that the survey did not involve much private information, each respondent answered 7 single choice questions, including one question to distinguish their identity and 6 questions on their evaluations and preferences regarding cloud storage and cloud security. The 6 questions were:

(1) How much value do you think your data in cloud storage are worth? (2) Do you care about the cloud provider's ability to access your data in cloud storage? (3) Do you put private or confidential data in cloud storage? (4) Which kind of cloud provider do you prefer to, a large IT enterprise or a small business? (5) As an individual user, would you consider a third-party secure service provider? (6) Assume that you are an enterprise user, how would you choose your cloud security project?

For questions (1) to (4), we did not provide respondents any additional information so as to make them select in accordance with their own experiences and tendencies. For question (5), we provided them with a TSP offering two secure cloud services: a "personal edition" at the price of \$20/month and a "professional edition" at the price of \$200/month. The personal edition provides basic data encryption and security audits while the professional edition offers functions like web firewalls and remote options to make their data more secure. For question (6), we simulated an environment in which the respondent is an entrepreneur who has decided to use an enterprise cloud service and is hesitating over choosing which of two kinds of security product to adopt: the cloud security service provided by the CSP, or a TSP's security product. Each service is priced at \$1000/month and the entrepreneur can benefit by a gross profit of \$5000/month by using the cloud platform. The respondents could also choose to adpot both or neither of the two services.

7.2 Results

Figure 3 shows the respondents' evaluations of their data value in the cloud in response to question (1), classified by the respondents' identity. According to the results of the questionnaire, 5 persons (15.625%) thought their data had no value, 8 persons (25%) thought their data value was in the range \$0 – \$100, 8 persons (25%) thought their data value was in the range \$ 100 – \$1000, 7 persons (21.875%) thought their data value was in the range \$1000 – \$10 000 while 4 persons (12.5%) thought their data value was over \$10 000.

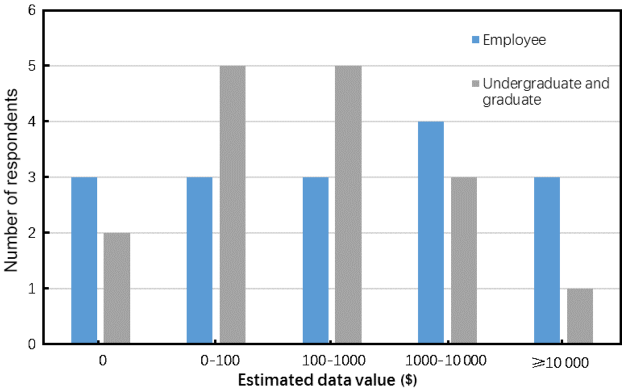

Fig. 3 Respondents' selection for their data value in cloud storage classified by their identity.

Figure 4 shows the respondents' attitudes towards cloud providers obtaining their data in response of question (2), classified by their selection of their data value in the cloud. In total, 6 persons (18.75%) did not care about their data being obtained by the CSP, 2 persons (6.25%) were happy to provide their data to the CSP in return for better service, 13 persons (40.625%) expressed concern about their data security and 11 persons (34.375%) were sufficiently concerned about data security that they would upload encrypted data to the cloud.

Figure 5 shows the distribution of how often the respondents uploaded their private or confidential data

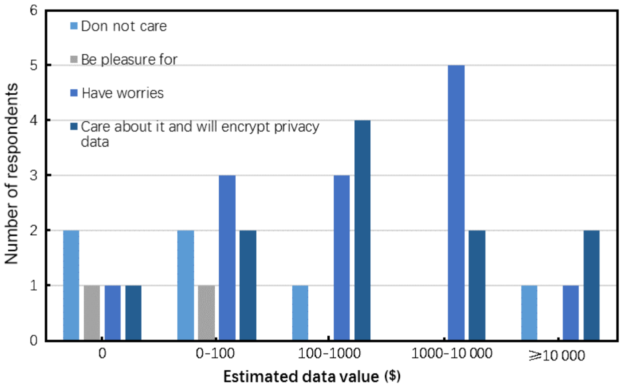

Fig. 4 Respondents' thoughts on cloud providers obtaining their data classified by their selection on data value.

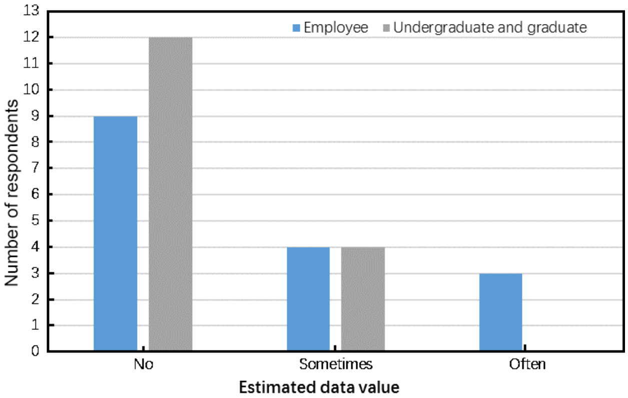

Fig. 5 Did respondents upload private or confidential data in cloud storage?

to cloud storage in response to question (3). In total, 21 persons (65.625%) said they would not store private or confidential data in the cloud, 8 persons (25%) said they sometimes did this and the remaining 3 persons (9.275%) said they often did so. For question (4), 22 persons (68.75%) preferred the cloud service of large enterprise, 2 persons (6.25%) preferred small businesses, while 8 persons (25%) did not trust any CSP.

For question (5), Fig. 6 shows respondents' choices on whether they engage a TSP classified by their data value in the cloud. In total, 14 persons (43.75%) chose not to use a TSP's service, 16 persons (50%) chose the personal edition, and 2 persons (6.25%) chose the professional edition.

For question (6), 9 persons (28.125%) chose to use the CSP's security service, 13 persons (40.625%) chose the TSP's service, 8 persons (25%) used both of them, and 2 persons (6.25%) chose not to use either.

7.3 Analysis and discussions

According to Fig. 3, we can see that majority (87.5%) respondents thought their data in cloud storage were worth less than $10 000. The t test between employee and student resulted in t=1.121 and p=0.136, one tailed, showing that users' identity is not relevant to their data value in the cloud. We can see that neither employees nor students often use the cloud to store high-value data. From Fig. 4, 25% persons did not care about their data being obtained by the CSP. Among the respondents who thought their data value in cloud was 0, a majority (60%) did not care about their data being obtained by the CSP. Using our one-user model can easily see that in this situation, users would not lose profits when their data is stolen, thus they would tend not to mind, which is consistent with the survey results. Among the respondents who thought their data value in cloud was $100 – $1000, this percentage decreases to 37.5%. In other classes of respondents, the people who cared about the data stealing became a majority. For high value data, the majority of users think in a mode that follows our model's assumption.

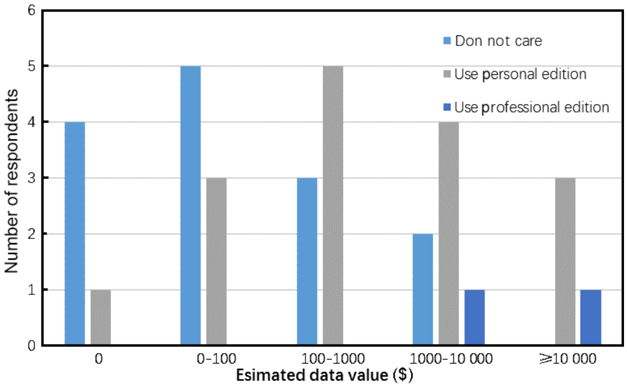

Fig. 6 Respondents' selection on third-party secure service classified by their selection on data value.

We can see from Fig. 5 that a majority of respondents (65.625%) would not store private or confidential data in the cloud, which means they were concerned about the security of outsourced data storage even though the data value they were storing on the cloud was not very high, which is consistent with the result of questions (1) and (2). The t test between employee and student resulted in $t = 1.86$ and $p = 0.04$, one tailed, showing that users' identity is relevant to their data privacy in the cloud. We can see from Fig. 5 that, compared to students, employees uploaded more private and confidential data to cloud storage.

A majority of respondents (68.75%) preferred cloud services from large enterprises, and only 6.25% trusted small businesses, which suggests that most individual users equate a CSP's trust degree with the reputation of the enterprise.

According to the result of question (5) shown in Fig. 6, most people preferred to use a TSP's security service. Moreover, the percentage of people who preferred to use a TSP increases as their data value increases. Especially, the respondents whose data were valued higher than $10 000 all chose to use the TSP. However, only 2 persons (6.25%) chose the professional edition product, both of whom held data in the cloud valued at greater than $1000. This is also consistent with our model because of the price of the professional edition service. When data is valued at less than $1000, using the professional edition with a price of $200/month would quickly lead to negative profits, thus the users would never choose it.

For the result of question (6), when the price of the security service provided by the CSP and the TSP are equal, most users preferred to use the TSP. This indicates that a majority of users hold the TSP to be more secure than the CSP, which can also be confirmed from the response to question (5). Because in this situation the cost of using the CSP and the TSP is the same, we can conclude through our analysis that the majority thought the TSP's trust degree to be higher than the CSP's. Otherwise, choosing the TSP means that users pay more without gaining additional security.

Through the analysis of the questionnaire, we can see our models work for a majority of individual cloud users and have potential for the security improvement of real systems. The majority of users did not store high value data in cloud storage, would not be willing to pay much for securing their data, and preferred cloud storage offerings from large enterprises. These conclusions can be guidance for companies to conduct their cloud business. Furthermore, for enterprise cloud users, who are more rational than individual users, our model provides more quantitative tools for them to make decision around adopting cloud computing.

8 Conclusion

In this paper, we constructed a series of cloud security models through game theory to assess the risk of the cloud service providers stealing users data. The models we constructed can assess the internal security hazards in the existing cloud systems and estimate whether a CSP or a TSP would behave honestly to user data.

When analyzing a game between users and service providers in a cloud system, if a Nash equilibrium strategy profile makes each SP who is chosen by a user would choose not to steal the user's data according to his utility function, then the cloud system has theoretic internal security. A semi-trustworthy TSP will offer users additional security as long as users trust it at least as highly as the CSP.

Our experimental results show our model and conclusion to be practical and to work for the majority of users who participated in our experiment. We believe our work has potential in real cloud environments and can serve as guidance for enterprises and individuals in providing and making use of cloud services.

References

[1] A. Reed, C. Rezek, P. Simmonds, eds., Security guidance for critical areas of focus in cloud computing v3.0, http://cloudsecurityalliance.org/guidance/, 2011.

[2] D. Cuschieri, Cloud encryption and key management considerations, Tech. report, RHUL–MA–2014–9, University of London, Royal Holloway, UK, 2014.

[3] International Organization for Standardization, *ISO 31000, Risk Management: Principles and Guidelines*. 2009.

[4] J. O. Fitó, M. Mácias, and J. Guitart, Toward business driven risk management for cloud computing, in *2010 International Conference on Network and Service Management (CNSM)*, 2010, pp. 238–241.

[5] E. Furuncu and I. Sogukpinar, Scalable risk assessment method for cloud computing using game theory (CCRAM), *Computer Standards & Interfaces*, vol. 38, pp. 44–50, 2015.

[6] U. Wazir, F. G. Khan, S. Shah, Service level agreement in cloud computing: A survey, *International Journal of Computer Science and Information Security*, vol. 14, no. 6, p. 324, 2016.

[7] J. Li, J.W. Li, and X. F. Chen, Identity-based encryption with outsourced revocation in cloud computing, *IEEE Transactions on Computers*, vol. 64, no. 2, pp. 425–437, 2015.

[8] X. Yi, F. Y. Rao, and E. Bertino, Privacy-preserving association rule mining in cloud computing, in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, 2015, pp. 439–450.

[9] Y. Yong, M. H. Au, and G. Ateniese, Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage, *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 767–778, 2017.

[10] P. Narwal, D. Kumar, and M. Sharma, A review of game-theoretic approaches for secure virtual machine resource allocation in cloud, in *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 2016.

[11] M. H. Manshaei, Q. Y. Zhu, and T. Alpcan, Game theory meets network security and privacy, *ACM Computing Surveys (CSUR)*, vol. 45, no. 3, p. 25, 2013.

[12] R. Anderson and T. Moore, The economics of information security, *Science*, vol. 314, no. 5799, pp. 610–613, 2006.

[13] L. J. Camp and S. Lewis, eds., *Economics of Information Security*. New York, NY, USA: Kluwer, 2006.

[14] R. Böhme and G. Schwartz, Modeling cyber-insurance: Towards a unifying framework, presented at Workshop on the Economics of Information Security (WEIS), Cambridge, MI, USA, 2010.

[15] J. Grossklags, N. Christin, and J. Chuang, Secure or insure?: A game-theoretic analysis of information security games, in *Proceedings of the 17th International Conference on World Wide Web*, 2008, pp. 209–218.

[16] J. Grosslags and B. Johnson, Uncertainty in the weakestlink security game, in *Game Theory for Networks, 2009. GameNets' 09. International Conference on*, 2009, pp. 673–682.

[17] J. Lou and Y. Vorobeychik, Equilibrium analysis of multi-defender security games, in *Proceedings of the Twenty-Fourth International Joint Conference on Artifical Intelligence (IJCAI)*, 2015, pp. 596–602.

[18] D. Ardagna, B. Panicucci, and M. Passacantando, A game theoretic formulation of the service provisioning problem in cloud systems, in *Proceedings of the 20th International Conference on World Wide Web*, 2011, pp. 177–186.

[19] E. Bertino and E. Ferrari, Secure and selective dissemination of XML documents, *ACM Transactions on Information and System Security (TISSEC)*, vol. 5, no. 3, pp. 290–331, 2002.

[20] M. Gerome and S. Dan, Controlling access to published data using cryptography, in *Proceedings of the 29^{th} International Conference on Very Large Data-bases*, 2003, pp. 898–909.

[21] S. D. Di Vimercati, S. De Capitani, and S. Foresti, Overencryption: Management of access control evolution

on outsourced data, in *Proceedings of the 33rd International Conference on Very Large Data-bases*, 2007, pp. 123–134.

[22] V. Goyal, O. Pandey, and A. Sahai, Attribute-based encryption for fine-grained access control of encrypted data, in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, 2006, pp. 89–98.

[23] A. Shamir, Identity-based cryptosystems and signature schemes, in *Workshop on the Theory and Application of Cryptographic Techniques*, 1984, pp. 47–53.

[24] G. J. Wang, Q. Liu, and J. Wu, Hierarchical attribute-based encryption for fine-grained access control in cloud storage services, in *Proceedings of the 17th ACM Conference on Computer and Communications Security*, 2010, pp. 735–737.

[25] M. Nabeel, N. Shang, J. Zage, and E. Bertino, Mask: A system for privacy-preserving policy-based access to published content, in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, 2010, pp. 1239–1242.

[26] M. Nabeel, N. Shang, and E. Bertino, Privacy preserving policy-based content sharing in public clouds, *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 11, pp. 2602–2614, 2013.

[27] M. Nabeel and E. Bertino, Privacy preserving delegated access control in public clouds, *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 9, pp. 2268–2280, 2014.

[28] N. K. Sharma and A. Joshi, Representing attribute based access control policies in owl, in *2016 IEEE Tenth International Conference on Semantic Computing(ICSC)*, 2016, pp. 333–336.

[29] A. Sangroya, S. Kumar, J. Dhok, and V. Varma, Towards analyzing data security risks in cloud computing environments, in *International Conference on Information Systems, Technology and Management*, 2010, pp. 255–265.

[30] B. S. Kaliski Jr and W. Pauley, Toward Risk assessment as a service in cloud environments, in *HotCloud'10 Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*, 2010, p. 13.

[31] M. Theharidou, N. Tsalis, and D. gritzalis, In cloud we trust: Risk-assessment-as-a-service, in *IFIP International Conference on Trust Management*, 2013, pp. 100–110.

[32] S. Drissi, H. Houmani, and H. Medromi, Survey: Risk assessment for cloud computing, *International Journal of Advanced Computer Science and Applications*, vol. 412, 2013.

[33] Z. Ismail, C. Kiennert, J. Leneutre, and L. Chen, Auditing a cloud provider's compliance with data backup requirements: A game theoretical analysis, *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1685–1699, 2016.

Yuzhao Wu received the BS degree from Tsinghua University, Beijing, China in 2011. He is currently pursuing the PhD degree in Institute for Interdisciplinary Information Sciences at Tsinghua University, Beijing, China. His research interests lies in the area of theoretical computer science, cryptography, and information security.

Yongqiang Lyu received the BS degree from Xidian University, Xi'an, China, in 2001, and the MS and PhD degrees in computer science from Tsinghua University, Beijing, China, in 2003 and 2006, respectively. He is currently an associate professor with the Research Institute of Information Technology, Tsinghua University. His research interest focuses on the hardware-software fusion architecture in emerging computing systems.

Yuanchun Shi received the BS, MS, and PhD degrees in computer science from Tsinghua University, Beijing, China in 1989, 1993, and 1999, respectively. She is a Changjiang Distinguished Professor with the Department of Computer Science, Tsinghua University. She was a Senior Visiting Scholar with MIT AI Lab during 2001-2002. She has authored and co-authored more than one hundred papers in *International Journal of Human-Computer Studies*, *IEEE Transactions on Parallel and Distributed Systems*, *IEEE Transactions on Knowledge and Data Engineering*, *ACM Transactions on Computer-Human Interaction*, *ACM Multimedia*, *ACM User Interface Software and Technology*, etc. Her research interests include human-computer interaction, pervasive computing, and multimedia communication. Dr. Shi had chaired several conferences including ACM Ubicomp2011. She serves as the Area Editor of *Pervasive and Mobile Computing* (Elsevier), an editor of the *Interacting With Computer* (Oxford University Press), and the Vice Editor-in-Chief of the Communications of China Computer Federation.