Peer Prediction Based Trustworthiness Evaluation and Trustworthy Service Rating in Social Networks

Article *in* IEEE Transactions on Information Forensics and Security · November 2018

DOI: 10.1109/TIFS.2018.2883000

6 authors, including:

Jun Du
Tsinghua University
39 PUBLICATIONS **362** CITATIONS

SEE PROFILE

Erol Gelenbe
Polish Academy of Sciences
758 PUBLICATIONS **17,871** CITATIONS

SEE PROFILE

Chunxiao Jiang
Tsinghua University
332 PUBLICATIONS **5,490** CITATIONS

SEE PROFILE

Peer Prediction Based Trustworthiness Evaluation and Trustworthy Service Rating in Social Networks

Jun Du, *Member, IEEE*, Erol Gelenbe, *Life Fellow, IEEE*, Chunxiao Jiang, *Senior Member, IEEE*, Haijun Zhang, *Senior Member, IEEE*, Yong Ren, *Senior Member, IEEE* and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—With the development of online applications based on the social network, many different approaches of service to achieve these applications have emerged. Users' reporting and sharing of their consumption experience or opinion can be utilized to rate the quality of different approaches of online services. How to ensure the authenticity of the users' reports and identify malicious ones with cheating reports become important issues to achieve an accurate service rating. In this paper, we provide a private-prior peer prediction mechanism based trustworthy service rating system with a data processing center (DPC), which requires users to report to it with their prior and posterior believes that their peer users will report a high quality opinion of the service. The DPC evaluates users' trustworthiness with their reports by applying the strictly proper scoring rule, and removes reports received from users with low trustworthiness from the service rating procedure. This peer prediction method is incentive compatible and able to motivate users to report honestly. In addition, to identify malicious users and bad-functioning/unreliable users with high error rate of quality judgement, an unreliability index is proposed in this paper to evaluate the uncertainty of reports. Reports with high unreliability values will also be excluded from the service rating system. By combining the trustworthiness and unreliability, malicious users will face a dilemma that they cannot receive a high trustworthiness and low unreliability at the same time when they report falsely. Simulation results indicate that the proposed peer prediction based trustworthy service rating can identify malicious and unreliable behaviours effectively, and motivate users to report truthfully. The relatively high service rating accuracy can be achieved by the proposed system.**

***Index Terms*—Service rating, trustworthiness, reliability, peer prediction, private prior, social networks.**

This work was supported in part by the National Natural Science Foundation China under Projects 61571300, 61428101, 61822104, 61471025, and 61771044, in part by the Young Elite Scientist Sponsorship Programs by CAST, in part by the new strategic industries development projects of Shenzhen City under Grant JCYJ20170816151922176, and in part by the Research Foundation of the Ministry of Education of China and China Mobile under Grant MCM20170108. (*Corresponding author: Chunxiao Jiang.*)

J. Du and Y. Ren are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, P. R. China (e-mail: blgdujun@gmail.com, reny@tsinghua.edu.cn).

E. Gelenbe is with the Electric and Electronic Engineering Department, Imperial College London, South Kensington Campus, London SW7 2AZ, UK (e-mail: e.gelenbe@imperial.ac.uk).

C. Jiang is with the Tsinghua Space Center, Tsinghua University, Beijing 100084, China, and also with the Key Laboratory of EDA, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China (e-mail: jchx@tsinghua.edu.cn).

H. Zhang is with the Beijing Advanced Innovation Center for Materials Genome Engineering, Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, Institute of Artificial Intelligence, University of Science and Technology Beijing, Beijing 100083, China (e-mail: haijunzhang@ieee.org).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

I. Introduction

Information communication and computation technologies have been developing rapidly in recent years. With the growing demands of big data and development of different applications, the emerging fifth generation (5G) mobile communication technology will be a multi-service and multi-technology integrated network, which can enhance the user experience by providing various intelligent and customized services [1]. Moreover, social networks have become important platforms for users to enjoy different kinds of online services. With the rapid development of Internet-based applications, different approaches to achieve these applications have emerged. Take e-commerce for an instance, in which users are allowed to use different online or mobile payment systems, such as PayPal, Google Wallet, Alipay and Apple Pay, to complete payments. In addition, for some file sharing applications, users can use different downloaders to download their favorite music, movies or other media files.

In order to provide accurate and useful suggestions to new users and help them to make choices, the use of service quality ratings for these different services has become an important method [2], [3], [4]. Concerning this problem, the feedback and evaluation from users who have experienced a service provide essential reference information for the service rating [5], [6], [7]. Meanwhile, social networks provide platforms that collect and share users' feedback, according to which the service rating can be provided through some data fusion mechanism. However, false and dishonest reports from malicious users can destroy the fairness and usefulness of such ratings. Therefore, it is rather necessary to introduce some trust assessment function to such systems and design an incentive mechanism to motivate users to output truthful feedback.

In this paper, we will establish a peer prediction based trustworthy service rating system for social networks. With peer prediction based decision, network functions of malicious behavior detection, trustworthiness and unreliability assessment can be achieved. Then the reliable and trustworthy service ratings can be obtained by the feedback from honest and reliable users. In this work, we assume that the service quality is an objective evaluation independent of users' subjective judgements. This assumption is reasonable for many service quality indicators, such as convenience of online payment methods and download speeds [8], [9].

A. Literature Review

Service ratings for different application systems have been active research topics over the past decades. Many service

evaluation systems have been developed for mobile social networks [10], multiple providers service systems [11] and many other kinds of web services [12], [13]. In [14], researchers designed the objective rating scores of products or services through an iterative rating algorithm. This rating mechanism entirely decoupled the credibility assessment of the evaluations from the ranking itself, which makes it very robust against collusion attacks as well as random and biased raters. A two-phase methodology was proposed in [15] for systematically evaluating the performance and availability of cluster-based Internet services. A service rating scheme that is robust against manipulations by malicious users and services was proposed in [16]. In [16], the service rating made by the target customer was predicted, based on which the system helped this customer to choose a suitable service. The authors of [17] proposed a user-service rating prediction approach for the recommender system by exploring social users' rating behaviors. In [17], the user's social relationships were considered in order to understand social users' rating behavior diffusions.

A social network is a platform that allows its users to obtain services and share their experiences [18]. Based on such feedback gathered, a data processing center (DPC) can provide quality ratings for different services, which can further give suggestions for new users. To ensure the accuracy of service ratings, the trustworthiness and reliability of the feedback from users need to be checked and ensured. Currently, trust and reputation management has become a challenge in many kinds of feedback and decision systems. Many trustworthiness evaluation mechanisms have been proposed for social networks [19], [20], wireless sensor networks [21], [22] and cloud-based service systems [23]. To motivate secondary users (SUs) in a multiple channel cognitive radio network to report truthfully, a Stackelberg game model was designed in [24], according to which trustworthy SUs gain transmission opportunities as rewards. A consumer feedback based service rating system was presented in [25] to evaluate the trustworthiness of a cloud service. In [25], a novel protocol was proposed to improve and ensure the credibility of trust feedback from consumers. In [26], a dynamic trust evaluation model was proposed to evaluate the user's reputation. The authors of [26] considered both users' preferences for different quality of service attributes and the impact of vicious ratings on trust evaluation. For rating the reputation of the service, different users' ratings were weighted dynamically according to their honesty assessment, and the influence of malicious ratings were thus effectively diluted.

Most of the local and global trustworthiness evaluation methods mentioned above are established by users' own current and/or past behaviors. Further, some researchers have considered relationship and interaction among users of a network for user trustworthiness assessment and prediction [27], [28], [29], [30], [31], although the incentive mechanisms for truthful information are not studied much. Originally applied in electronic commerce, common-prior peer prediction with a strictly proper scoring rule [32], [33] was proposed for truthful feedback from users in [34]. To be specific, *Peer Prediction* refers to a scheme using one user's report to update or predict a probability distribution for the report of someone else, whom we refer to as the "peer". The former user is then scored not on a comparison between the likelihood assigned to the peer's possible ratings and the peer's actual rating. Moreover, in the common-prior peer prediction mechanism, the prior probability of the product type or service quality is commonly held, conditional on which, the probability distribution of user's received product type or service quality is also common knowledge. Relaxing the assumption of common-prior, the authors of [35] modified the classical peer prediction method such that only users' subjective and private opinions were needed, and this trustworthiness evaluation mechanism is known as private-prior peer prediction. Both of these two peer-prediction methods estimated the trustworthiness using strictly proper scoring rules, which can provided incentives for truthful reporting. The peer prediction mechanism can be applied efficiently in the scenario where the prior knowledge is subjective and private to each users. For instance, peer prediction has been used in wireless sensor networks [36], [37], cognitive radio networks [38], [39], social and online systems [40], [41] and many other kinds of crowd-sourcing systems [42], [43] to collect truthful reports from users, and has been considered as an effective solution to elicit trustworthy feedback. In this paper, we propose a service rating system for social network based services according to honest users with high trustworthiness. Private-prior peer prediction is introduced to evaluate users' trustworthiness and motivate users to provide truthful feedback.

B. Contributions and Organization

The main contributions of this paper can be summarized as follows.

- We introduce private-prior peer prediction in the service rating system of social networks. The user trustworthiness obtained through certain strictly proper scoring rules is formulated to motivate users to report trustfully. We analyze the incentive compatibility of the basic peer prediction mechanism with respect to the false alarm and missed detection probabilities of judgement and report.
- We propose an unreliability index to eliminate unreliable reports from the service rating system. By applying the unreliability index, malicious users are confronted with a dilemma that they cannot get a high trustworthiness and a low unreliability at the same time when they provide a false report. However, the best choice of honest users is still reporting truthfully even for poorly functioning ones with high error rates of judgement.
- Based on the proposed user trustworthiness and unreliability index, we design a service rating framework. In this framework, trustworthiness is used to evaluate the possibility of whether the subject user's report is dishonest and the user is a malicious one. On the other hand, the unreliability index is introduced to determine whether the reports are reliable, but does not consider the type of the users, i.e., honest or malicious. By removing the feedback reports with high unreliability and reports received from users with low trustworthiness, from the final rating procedure, an accurate and trustworthy service rating can be achieved.

The rest of this paper is organized as follows. In Section II, the system model is described. The private-prior peer prediction based user trustworthiness evaluation for motivating truthful reports is proposed in Section III. Then we analyze the reliability of users' reports and design the service rating system in Section IV. Simulations are presented in Section V, and conclusions are drawn in Section VI.

II. Mathematical Model for Service Rating Based on User Report Fusion

With the boom of online applications based on social networks, different services to support these applications have emerged [44], [45]. As mentioned previously, users are allowed to select different online payment methods to complete their online purchases, or download their favorite music and movies by downloaders those they think are faster and more reliable. To rate the quality of different services, users are required to report and share their consumption experiences or opinions to the social application platform, which can use this valuable feedback for service rating and helping new users to judge whether the applications can provide high quality services. In our work, the quality of the services is considered as an objective evaluation independent of users' subjective judgements. For instance, different users tend to have the same opinion about whether a payment system is convenient or a downloader has a high download speed. Such a social rating system is different from systems such as movie review, in which users' subjective opinions and standards may vary considerably between individuals.

In this case, users' truthful feedback of a service is important for achieving an accurate rating of this service approach's quality and providing helpful suggestions to new users. However, some malicious users in social networks provide untruthful evaluations of the service quality for some purposes. On the one hand, malicious users report to the service rating system that the object approach of service is high in quality when it gives a bad service performance to improve its competitiveness. On the other hand, malicious users report a low service-quality evaluation to lower the rating of the object approach of service, which will encourage new users not to select it. These malicious behaviours undermine the fairness of the service rating and provide unreliable suggestions to new users. Therefore, it is important to make sure that the feedback from users is truthful.

In this work, we design a mechanism to provide incentives for truthful opinions of users. Moreover, we define a trustworthiness management method to identify malicious users, excluding whose untruthful feedback, the service rating with high accuracy can be made.

A. System Model

Consider a population of N users distributed over a social network with a service platform, which can provide different approaches of this service. Quality Q of the service is a binary rating, which is considered as a random variable represented by $\{l, h\}$ referring to the low quality and high quality, respectively. As mentioned previously, this quality is an objective fact. In other words, after experiencing the service, different honest users tend to give the same evaluation or opinion independent of their individual subjective standards. As shown in Fig. 1, each user i ($i = 1, 2, \cdots, N$) accepts the service m, and then makes a binary opinion of the service quality denoted by $S_i = s_i \in \{l, h\}$. Meanwhile, users are allowed to provide some required QoS reports to the cloud, and these reports will be processed by the DPC. For instance, the opinion report denoted by $x_i \in \{0, 1\}$ is generated by applying a report strategy $r_i : S_i \to \{0, 1\}$. User i will report $x_i = 1$ when $S_i = h$ (or $x_i = 0$ when $S_i = l$) to the cloud if he/she is honest. We assume that x_i is the semi-public information published to the social service-evaluation platform by the cloud, and can be observed by the DPC and other users over the social network and having the friendship with user i. In addition, S_i is the private or local information only known by user i, and other users and even the cloud cannot get it.

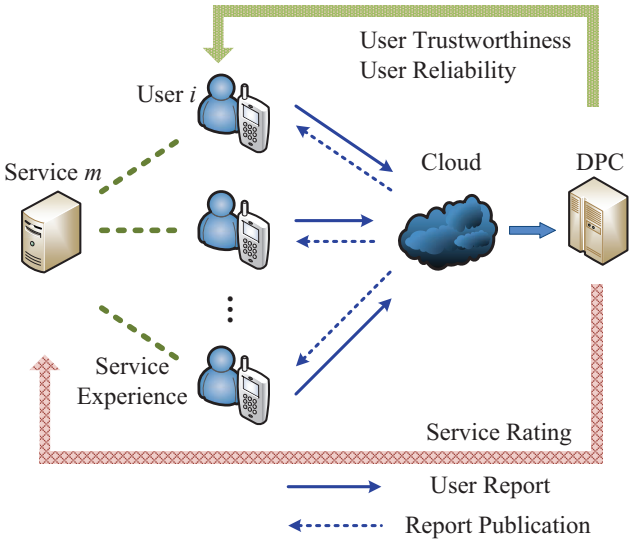


Fig. 1. Peer prediction based service rating and user trustworthiness management system.

B. Service Rating Based on User Report Fusion

Define the false alarm of the judgement as that the service is misjudged as a low quality while it is a high quality in fact, and the user i's false alarm probability of judgement is denoted by $P_{fa,i} = P(S_i = l \,|Q = h)$. In order to simplify the expression, let P_{fa} to denote the false alarm probability of judgement. On the contrary, define the missed detection probability of judgement as $P_{md,i} = P(S_i = h \,|Q = l)$. Similarly, we use P_{md} to denote the missed detection probability of judgement for a simpler expression. As mentioned previously, the quality is an objective fact, which leads that both honest and malicious users trend to make the similar and accurate judgement for it. So we assume that $P_{fa} < 0.5$ and $P_{md} < 0.5$ hold for all users in the social network.

On the other hand, we consider the false alarm of the report as that a user reports a low quality evaluation to the cloud when the quality of service is high, and the user i's false alarm probability of report is $P_{f,i} = P(x_i = 0 \,|Q = h)$. The missed

detection probability of report is $P_{m,i} = P(x_i = 1 | Q = l)$. In addition, the simplified expression of the two probabilities of report above are P_f and P_m. We consider that the user type are represented by $\theta_i \in \{0, 1\}$, i.e., $\theta_i = 0$ if i is an honest user, and $\theta_i = 1$ if user i is the malicious otherwise. Assume that if user i is malicious, his/her false alarm cheating rate is $P^c_{f,i} \in [0, 1]$, and the missed detection cheating rate is $P^c_{m,i} \in [0, 1]$. We assume that the honest users always report their real judgement of the service quality, no matter whether his/her judgement is accurate. Then for each user i, we have

$$P_{f,i} = \begin{cases} (1 - P_{fa,i}) P^c_{f,i} + P_{fa,i} \left(1 - P^c_{f,i}\right), & \theta_i = 1; \\ P_{fa,i}, & \theta_i = 0; \end{cases} \tag{1}$$

$$P_{m,i} = \begin{cases} (1 - P_{md,i}) P^c_{m,i} + P_{md,i} \left(1 - P^c_{m,i}\right), & \theta_i = 1; \\ P_{md,i}, & \theta_i = 0. \end{cases} \tag{2}$$

As shown in Fig. 1, based on the users' reports received by the cloud, the DPC can obtain trustworthiness T_i of each user and make the decision of the service rating by applying the following rule:

$$R = \sum_{i \in T} x_i \begin{cases} < n, & \text{the DPC rates the service } Q = l; \\ \geq n, & \text{the DPC rates the service } Q = h, \end{cases} \tag{3}$$

where n is the threshold of service rating. In (3), $T = \{i \,|T_i \geq t\}$ is the set of honest users with high trustworthiness T_i, which is determined by the threshold of trustworthiness t. A simple decision making rule is that the DPC rates the service as high, i.e., $Q = h$, only if more than half of trust users report the service's quality is high, i.e., $n = |T| / 2$.

III. Peer Prediction for User Trustworthiness

In this section, we will introduce the private-prior peer prediction method, which enable to encourage users to provide rating reports truthfully. With some certain strictly proper scoring rules to estimate the users' trustworthiness, the mechanism can identify malicious users those with low trustworthiness. Then users are motivated to report truthfully in order to obtain high trustworthiness and avoid being considered as the malicious.

A. Private-Prior Peer Prediction Mechanism

Private-prior peer prediction is an incentive compatible mechanism originally proposed to motivate agents to report their private prior and posterior signal belief on electronic commerce [35]. In the basic private-prior peer prediction mechanism, each agent i coupled with his/her peer agent $j = i + 1$ is required to report his/her prior and posterior signal belief of the state before and after observe the signal, respectively. According to the two reports, the agent i's score can be calculated by a *strictly proper scoring rule*, which will be introduced in the later part of this section.

1) Prior belief reports to the cloud: In the system established in this work, any two users accepting the same service can be considered as a pair of peers, which establishes a kind of friendship and topology of all users. For rating the quality of service m, we consider that each user i has one peer user j selected randomly from other users who have accepted and will still accept the same service m as i. Before experiencing the service, user i is required to report his/her prior belief $y_{ij} \in [0, 1]$, or called *information report*, to the cloud that his/her peer user j will report a high quality signal, i.e., $x_j = 1$. Then y_{ij} can be given by

$$\begin{aligned} y_{ij} &= P_i(x_j = 1) \\ &= P_i(x_j = 1 | Q = h) P_i(Q = h) \\ &\quad + P_i(x_j = 1 | Q = l) P_i(Q = l) \\ &\triangleq P(x_j = 1 | S_i = h) P(S_i = h) \\ &\quad + P(x_j = 1 | S_i = l) P(S_i = l). \end{aligned} \tag{4}$$

In (4), $P_i(x_j = 1 | Q = h)$ and $P_i(x_j = 1 | Q = l)$ can be obtained by the previous report x_j of user j released among the network. $P_i(x_j = 1 | Q = h)$ represents the probability that user j gives a report of "high quality" evaluation for the service when user i makes a high quality judgement to the same service, i.e., $S_i = h$. This judgement is a private and local information only known by user i, and the prior belief $P_i(Q = h)$ is user i's subjective prior of the service quality and is identical to $P(S_i = h)$. Similarly, $P_i(Q = l)$ is equal to $P(S_i = l)$. Therefore, we can get the second equivalence relation in (4) established by $\triangleq$.

2) Posterior belief reports to the cloud: After experiencing the service, user i makes his/her own opinion of the service quality $S_i = s_i$, and then sends the posterior belief, or called *prediction report* to the cloud, denoted by $y'_{ij}(s_i) \in [0, 1]$, that the peer user j will report of a high quality evaluation for the service. Then y'_{ij} can be expressed as

$$\begin{aligned} y'_{ij}(s_i) &= P_i(x_j = 1 | S_i = s_i) \\ &= P(x_j = 1 | Q = h) P(Q = h | S_i = s_i) \\ &\quad + P(x_j = 1 | Q = l) P(Q = l | S_i = s_i). \end{aligned} \tag{5}$$

Similar to the previous analysis, $y'_{ij}(s_i)$ can be decomposed into two conditions as follows.

$$y'_{ij}(l) = \frac{\varphi_1 (1 - P_{f,j}) + \varphi_2 P_{m,j}}{\varphi_1 + \varphi_2}, \tag{6a}$$

$$y'_{ij}(h) = \frac{\varphi_3 (1 - P_{f,j}) + \varphi_4 P_{m,j}}{\varphi_3 + \varphi_4}, \tag{6b}$$

where

$$\varphi_1 = P_{fa,i} P(Q = h), \varphi_2 = (1 - P_{md,i}) P(Q = l), \tag{7a}$$

$$\varphi_3 = (1 - P_{fa,i}) P(Q = h), \varphi_4 = P_{md,i} P(Q = l). \tag{7b}$$

As defined previously, y_{ij} is the user i's prior judgement that $x_j = 1$ before user i experiences the service. After user i experiencing the service and sensing that $s_i = h$, it is reasonable for user i to make the judgement that $x_j = 1$ with a larger probability, i.e., $y'_{ij}(h) > y_{ij}$, which means that i's prior belief $x_j = 1$ will be "boosted". On the contrary, $y_{ij} > y'_{ij}(l)$ if user i receives a low-quality service. However, when there are malicious users providing untrustful evaluations of the service, the relation of inequality above cannot always satisfied. Lemma 1 provides the sufficient conditions which can ensure $y'_{ij}(h) > y_{ij} > y'_{ij}(l)$.

Lemma 1. *In the private-prior peer prediction mechanism, for each user i with prior and posterior belief reports y_{ij} and y'_{ij} of user j, it holds that $y'_{ij}(h) > y_{ij} > y'_{ij}(l)$ if all users satisfy that $P_{fa} + P_{md} < 1$ and $P_f + P_m < 1$.*

Proof: See Appendix.

Remarks: As been assumed that $P_{fa} < 0.5$ and $P_{md} < 0.5$, condition $P_{fa}+P_{md} < 1$ always holds for all users. According to (1) and (2), for honest user i, i.e., $\theta_i = 0$, we have $P_{f,i} + P_{m,i} < 1$. On the other hand, for dishonest user i ($\theta_i = 1$), whether $P_{f,i}+P_{m,i} < 1$ can hold depends on his/her false alarm cheating rate $P^c_{f,i}$ and missed detection cheating rate $P^c_{m,i}$. Notice that outright malicious users with relatively high $P^c_f > 0.5$ and/or P^c_m will have high $P_f > 0.5$ and/or $P_m > 0.5$, respectively. Users with both/either of the two cheating behaviors above can be identified easily according to their former reports with high error report probability. If the rating system removes reports of users having high former P_f and/or P_m, these malicious reports will not make sense when the system updates the rating of the service. Consequently, to achieve a continuous trick, malicious users need to manage their P^c_f and P^c_m to disguise themselves as trustful ones sometimes to make sure $P_f < 0.5$ and $P_m < 0.5$. So in our work, we analyze the peer prediction mechanism under the conditions of $P_f < 0.5$ and $P_m < 0.5$. Therefore, the condition of $P_f + P_m < 1$ in Lemma 1 is reasonable, and in this case, inequality $y'_{ij}(h) > y_{ij} > y'_{ij}(l)$ can be always satisfied.

3) Inferred opinion reports: Instead of reporting the private evaluation of the service quality S_i or x_i, user i sends his/her prior and posterior probability of belief that peer user j gives report $x_j = 1$. We notice that both report x_i and x_j are not provided directly by the relative user. In basic private-prior peer prediction, user i only sends reports y_{ij} and $y'_{ij}(s_i)$ to the cloud, according to which the DPC infers opinion report x_i and publishes it to the social service-evaluation platform. Inferred opinion report x_i is generated by the following rule:

$$x_i = x\left(y_{ij}, y'_{ij}\right) = \begin{cases} 1, & y'_{ij} > y_{ij}, \\ 0, & y'_{ij} < y_{ij}. \end{cases} \tag{8}$$

Remarks: According to Lemma 1, it holds that $y'_{ij}(h) > y_{ij} > y'_{ij}(l)$ when both user i and j satisfy $P_{fa} + P_{md} < 1$ and $P_f + P_m < 1$. In other words, when user i makes a high-quality judgement of the service after experiencing it ($S_i = h$), inequality $y'_{ij}(h) > y_{ij}$ always holds. Then applying (8), the DPC infers the opinion report as $x_i = 1$ because $y'_{ij} > y_{ij}$. So this inferred report $x_i = 1$ is consistent with user i's real judgement $S_i = h$. Symmetrically, when $S_i = l$, (8) can also derive the truthful opinion report $x_i = 0$. Therefore, the rule formulated by (8) can truthfully reflect the judgement when the user is honest, under the conditions of $P_{fa}+P_{md} < 1$ and $P_f + P_m < 1$.

4) User trustworthiness: Based on reports y_{ij} and $y'_{ij}(s_i)$, the DPC calculates user i's trustworthiness through a certain scoring rule. Users with low trustworthiness are classified as the malicious, and their reports will be unconsidered in the service rating system. Next, we first introduce the *strictly proper scoring rule*, which can motivate users to provide truthful reports y_{ij} and $y'_{ij}(s_i)$. The strictly proper scoring rule can be defined as Definition 1.

Definition 1. ***Strictly Proper Scoring Rule*** *[35]: A binary scoring rule is proper if it leads to an agent maximizing his/her score by truthfully providing his/her report $y \in [0, 1]$, and is strictly proper if an agent can maximize his/her score if and only if providing his/her report truthfully.*

The binary logarithmic and quadratic scoring rules shown as (9) and (10), respectively, are strictly proper, which has been proved in [32].

1) The binary logarithmic scoring rule:

$$R_l(y, \omega = 1) = \ln y \tag{9a}$$

$$R_l(y, \omega = 0) = \ln(1 - y). \tag{9b}$$

2) The binary quadratic scoring rule:

$$R_q(y, \omega = 1) = 2y - y^2, \tag{10a}$$

$$R_q(y, \omega = 0) = 1 - y^2. \tag{10b}$$

In (9) and (10), $\omega \in \{0, 1\}$ indicates a binary report.

We define the trustworthiness of user i as a function of y_{ij}, y'_{ij} and x_j:

$$T_i = \alpha R(y_{ij}, x_j) + (1 - \alpha) R\left(y'_{ij}, x_j\right) + \beta, \tag{11}$$

where $R(y, \omega)$ is a strictly proper scoring rule, $\alpha \in [0, 1]$ is the parameter weighting the importance of the prior and posterior belief. In addition, the trustworthiness will be cumulative as the service and scoring process continues. A negative trustworthiness can be a reflection of either monetary punishment or the limitation of report providing for the corresponding user, and the negative benefits will be transferred as positive benefits to the users as rewards for their honor and accurate reports. Therefore, to keep the budget balanced, β is given by

$$\beta = -\frac{1}{N} \sum_{k=1}^{N} \left[\alpha R(y_{kj}, x_j) + (1 - \alpha) R\left(y'_{kj}, x_j\right)\right]. \tag{12}$$

In (11), y_{ij} and y'_{ij} are the reports from user i before and after he/she makes judgement $S_i = s_i$ for the object service approach, respectively, and x_j is the user j's implicit opinion report inferred by the DPC according to user j's reports.

In addition, according to the analysis above, one can notice that the trustworthiness of user i is determined on user j's inferred opinion report x_j, user i's prior belief report y_{ij} and posterior belief report y'_{ij}. In other words, one user's trustworthiness is irrelevant to reports or inferred reports of the other users in the system. Therefore, the cooperative cheating of malicious users will have little effect on the evaluation of users' trustworthiness, which is defined by (11).

B. Incentive Compatibility

As proved in [35], prior belief report y_{ij} and posterior belief report $y'_{ij}(s_i)$ given by user i are temporal separated, which results from that they happen before and after making judgement $S_i = s_i$. Therefore, y_{ij} and $y'_{ij}(s_i)$ are independent

and then we have

$$\begin{aligned} & E\left[T_i\right] \\ &= E\left[\alpha R\left(y_{ij}, x_j\right)\right] + E\left[(1-\alpha) R\left(y'_{ij}, x_j\right)\right] + E\left[\beta\right] \\ &= \alpha\left(1-\frac{1}{N}\right) E\left[R\left(y_{ij}, x_j\right)\right] \\ &\quad + (1-\alpha)\left(1-\frac{1}{N}\right) E\left[R\left(y'_{ij}, x_j\right) | S_i = s_i\right] \\ &\quad - \frac{1}{N} \sum_{k=1, k\neq i}^{N} \left[\alpha R\left(y_{kj}, x_j\right) + (1-\alpha) R\left(y'_{kj}, x_j\right)\right], \end{aligned} \tag{13}$$

where both $\alpha\left(1-\frac{1}{N}\right)R(y_{ij}, x_j)$ and $(1-\alpha)\left(1-\frac{1}{N}\right)R\left(y'_{ij}, x_j\right)$ are still strictly proper [34].

1) Binary logarithmic scoring rule: We first apply the binary logarithmic scoring rule. Let $p_1 = P\left(x_j = 1\right)$ and $p_2 = P\left(x_j = 1 | S_i = s_i\right)$, and then we have

$$\begin{aligned} E\left[T_i\right] &= \alpha\left(1-\frac{1}{N}\right)\left[p_1 \ln y_{ij} + (1-p_1)\ln\left(1-y_{ij}\right)\right] \\ &= \alpha\left(1-\frac{1}{N}\right)\left[p_1 \ln y'_{ij} + (1-p_1)\ln\left(1-y'_{ij}\right)\right] \\ &\quad - \frac{1}{N}\sum_{k=1, k\neq i}^{N}\left[\alpha R\left(y_{kj}, x_j\right) + (1-\alpha) R\left(y'_{kj}, x_j\right)\right]. \end{aligned} \tag{14}$$

Take the partial derivatives with respect to y_{ij} and y'_{ij}:

$$\frac{\partial E\left[T_i\right]}{\partial y_{ij}} = \alpha\left(1-\frac{1}{N}\right)\frac{p_1 - y_{ij}}{y_{ij}\left(1-y_{ij}\right)} = 0, \tag{15a}$$

$$\frac{\partial E\left[T_i\right]}{\partial y'_{ij}} = \alpha\left(1-\frac{1}{N}\right)\frac{p_1 - y'_{ij}}{y'_{ij}\left(1-y'_{ij}\right)} = 0. \tag{15b}$$

Therefore we get the optimal values as

$$\hat{y}_{ij} = p_1 = P\left(x_j = 1\right), \tag{16a}$$

$$\hat{y}'_{ij} = p_2 = P\left(x_j = 1 | S_i = s_i\right). \tag{16b}$$

Then take the second partial derivatives with respect to y_{ij} and y'_{ij}, and let $y_{ij} = \hat{y}_{ij}$ and $y'_{ij} = \hat{y}'_{ij}$, then we have

$$\frac{\partial E^2\left[T_i\right]}{\partial y_{ij}^2}\Big|_{y_{ij}=\hat{y}_{ij}} = \alpha\left(1-\frac{1}{N}\right)\frac{y_{ij}\left(y_{ij}-1\right)}{y_{ij}^2\left(1-y_{ij}\right)^2} < 0, \tag{17a}$$

$$\frac{\partial E^2\left[T_i\right]}{\partial {y'_{ij}}^2}\Big|_{y'_{ij}=\hat{y}'_{ij}} = \alpha\left(1-\frac{1}{N}\right)\frac{y'_{ij}\left(y'_{ij}-1\right)}{{y'_{ij}}^2\left(1-y'_{ij}\right)^2} < 0. \tag{17b}$$

Therefore, the maximum of $E\left[T_i\right]$ can be achieved when $y_{ij} = p_1$ and $y'_{ij} = p_2$, which means that user i can receive the maximum trustworthiness if and only if he/she reports both y_{ij} and y'_{ij} truthfully.

2) Binary quadratic scoring rule: Next, we employ the binary quadratic scoring rule shown as (10). Thus we have

$$\begin{aligned} E\left[T_i\right] &= \alpha\left(1-\frac{1}{N}\right)\left[p_1\left(2y_{ij}-y_{ij}^2\right)+(1-p_1)\left(1-y_{ij}^2\right)\right] \\ &+(1-\alpha)\left(1-\frac{1}{N}\right)\left[p_2\left(2y'_{ij}-{y'_{ij}}^2\right)+(1-p_2)\left(1-{y'_{ij}}^2\right)\right] \\ &-\frac{1}{N}\sum_{k=1, k\neq i}^{N}\left[\alpha R\left(y_{kj}, x_j\right)+(1-\alpha) R\left(y'_{kj}, x_j\right)\right]. \end{aligned} \tag{18}$$

Take the partial derivatives with respect to y_{ij} and ${y'}_{ij}$, and set them to zero, we get the same optimal values as (16a) and (16b). Then take the second partial derivatives, the following inequality

$$\frac{\partial E^2\left[T_i\right]}{\partial y_{ij}^2} = \frac{\partial E^2\left[T_i\right]}{\partial {y'}_{ij}^2} = -2\alpha\left(1-\frac{1}{N}\right) < 0 \tag{19}$$

can be always satisfied.

Remarks: Noticing that $\partial^2 E\left[T_i\right]/\partial y_{ij}^2 < 0$ and $\partial^2 E\left[T_i\right]/\partial {y'_{ij}}^2 < 0$ will always be satisfied no matter whether the binary logarithmic or quadratic scoring rule is applied, the maximum of $E\left[T_i\right]$ can be reached when satisfying both (16a) and (16b). In other words, user i can receive the maximum trustworthiness if and only if he/she provides both y_{ij} and y'_{ij} truthfully, as mentioned previously. Assume that the cooperative cheating exists, which means that malicious users can contact with each other and manage the malicious behaviour. According to Definition 1, user i will obtain a lower score by reporting untruthfully than truthfully when his/her peer user j is a malicious one. For example, user i experiences a high-quality service and it means that his/her honest reports satisfy $y'_{ij} > y_{ij}$. However, because of user j's dishonest implicit opinion $x_j = 0$, user i will obtain a higher score if he/she gives a lower $y'_{ij} < y_{ij}$ instead of reporting truthfully, according to the binary logarithmic or quadratic scoring rule formulated as (9b) and (10b), respectively. To make sure that the honest users are predominant even when the cooperative cheating happens in the social network, we assume that the number of malicious users is less than the half of the total. Based on this assumption, the users with accurate information reports and prediction reports will always receive higher trustworthiness in a long term; meanwhile, the malicious users will be punished by a loss of trustworthiness every time they announce dishonest reports resulting in cheating opinion reports.

IV. User Trustworthiness and Unreliability based Service Rating

A. Unreliability of User Report

In private-prior peer prediction, all users are required to report their prior belief that their peer users will report a high evaluation for the service before experiencing the service $y_{ij} = P_i\left(x_j = 1\right)$. This report can be obtained by the past reports x_j inferred by the DPC and published by the cloud, which means that past reports x_j are accessible for i's other friends in the social network, the cloud and DPC. Therefore, it is difficult to fabricate information report y_{ij} for malicious users. To achieve cheating, malicious user i needs to manage his/her information and prediction report according to (8), i.e., $y'_{ij} = y_{ij} + \varepsilon$ ($\varepsilon > 0$) with probability $P^c_{m,i}$ when the service quality is low ($Q = l$), and $y'_{ij} = y_{ij} - \varepsilon$ with probability $P^c_{f,i}$ when the service quality is high ($Q = h$). Meanwhile, malicious users have to set ε as small as possible to avoid being punished by much loss of score and trustworthiness when their peer users are honest ones. In addition, we can notice that the false-alarm report and missed-detection report do not only result from the wrong judgements of honest users, but also due to

the dishonest users' cheating behaviours, according to (1) and (2). Both of the situations above are considered as unreliable behaviours which need to be identified and removed from the final service rating. Therefore, it is necessary to set a threshold to limit the minimum gap between y_{ij} and y'_{ij}.

Next, we analyze the influence of false-alarm judgement and missed-detection judgement on the scoring. Taking the derivative of (6a) and (6b) both with respect to $P_{fa,i}$ and $P_{md,i}$, we can calculate to get

$$\frac{\partial y'_{ij}(l)}{\partial P_{fa,i}} = \Phi_1 (1 - P_{md,i})(1 - P_{f,j} - P_{m,j}), \tag{20a}$$

$$\frac{\partial y'_{ij}(l)}{\partial P_{md,i}} = \Phi_1 P_{fa,i}(1 - P_{f,j} - P_{m,j}), \tag{20b}$$

$$\frac{\partial y'_{ij}(h)}{\partial P_{fa,i}} = -\Phi_2 P_{md,i}(1 - P_{f,j} - P_{m,j}), \tag{20c}$$

$$\frac{\partial y'_{ij}(h)}{\partial P_{md,i}} = -\Phi_2 (1 - P_{fa,i})(1 - P_{f,j} - P_{m,j}), \tag{20d}$$

where $\Phi_1 = P(Q=h)P(Q=l)/(\varphi_1+\varphi_2)^2$, $\Phi_2 = P(Q=h)P(Q=l)/(\varphi_3+\varphi_4)^2$. Based on the previous assumptions of $P_{fa,i} < 0.5$, $P_{md,i} < 0.5$ and $P_{f,j} + P_{m,j} < 1$, we have

$$\frac{\partial y'_{ij}(l)}{\partial P_{fa,i}} > \frac{\partial y'_{ij}(l)}{\partial P_{md,i}} > 0,\ \frac{\partial y'_{ij}(h)}{\partial P_{md,i}} < \frac{\partial y'_{ij}(h)}{\partial P_{fa,i}} < 0. \tag{21}$$

So under both of situations $Q = h$ and $Q = l$, the score of user i goes down with the increasing $P_{fa,i}$ and $P_{md,i}$ when user j reports truthfully, according to (9a)/(10a) and (9b)/(10b), respectively. In other words, for fixed P_f, P_j and $P(Q = h)$, the honest users with high judgement accuracy will receive higher scores and trustworthiness, compared to those honest users with high judgement error rates and malicious users reporting their prediction inversely and conservatively to give wrong reports and minimize the loss of scores. In the service rating system, neither implicit opinion reports of malicious users nor honest users with low judgement accuracy should be considered. To identify the two kinds of unreliable behaviour, we define an unreliability index to indicate the unreliability of user i by his/her prior belief report y_{ij} and posterior belief report y'_{ij} as follows.

$$\rho_i = \begin{cases} \frac{|y'_{ij} - P_{m,j}| P(Q=l)}{|y'_{ij} - (1-P_{f,j})| P(Q=h)}, & y'_{ij} < y_{ij}, \\ \frac{|y'_{ij} - (1-P_{f,j})| P(Q=h)}{|y'_{ij} - P_{m,j}| P(Q=l)}, & y'_{ij} > y_{ij}. \end{cases} \tag{22}$$

Remarks: In (22), the first situation $y'_{ij} < y_{ij}$ indicates that the more report y'_{ij} is closed to $P\{x_j = 1 | Q = l\}$ when the service quality is low and farther away from $P\{x_j = 1 | Q = h\}$ when the service quality is high, the more reliable y'_{ij} is. Meanwhile, for $y'_{ij} > y_{ij}$, when report y'_{ij} is closed to $P\{x_j = 1 | Q = h\}$ and far away from $P\{x_j = 1 | Q = l\}$, this report can be considered reliable. In addition, according to (21), $y'_{ij}(l)$ increases with growing $P_{fa,i}$ and $P_{md,i}$, and is more sensitive to $P_{fa,i}$ than $P_{md,i}$; $y'_{ij}(h)$ decreases with growing $P_{fa,i}$ and $P_{md,i}$, and is more sensitive to $P_{md,i}$ than $P_{fa,i}$. With assumption $P_{fa,i}, P_{md,i} \in [0, 1]$, we can get that $P_{m,j} < y'_{ij}(l), y'_{ij}(h) < 1 - P_{f,j}$, thus the definition of unreliability shown in (22) can be rewritten as

$$\rho_i = \begin{cases} \frac{[y'_{ij} - P_{m,j}] P(Q=l)}{[(1-P_{f,j}) - y'_{ij}] P(Q=h)}, & y'_{ij} < y_{ij}, \\ \frac{[(1-P_{f,j}) - y'_{ij}] P(Q=h)}{[y'_{ij} - P_{m,j}] P(Q=l)}, & y'_{ij} > y_{ij}. \end{cases} \tag{23}$$

To calculate the unreliability of users' reports, the DPC needs to observe the report error rates P_f and P_m of each user based on the historical reports and service rating results. In addition, we assume that the service quality, denoted by $P(Q = l)$ and $P(Q = h)$, can also be obtained according to a long time scale and relatively stable historical rating results of services. Such assumptions are feasible and reasonable, considering that most current service-based application systems have the ability to provide such information. By utilizing (23), the users with high unreliability ρ are considered to be uncertainty ones who might be honest users with high error judgement rate or malicious users. Reports from these users are not reliable for the DPC to rate the quality of service. Consequently, the DPC needs to set a threshold ρ_{thr}, and reports from the users with unreliability exceeding ρ_{thr} will be removed from the service rating procedure. The threshold can be designed by the typical error rates of honest users with relatively high judgement accuracy.

Next, we describe the validity of the user unreliability defined in (23). Take situation $Q = h$ for instance, malicious user i has to give the prediction report $y'_{ij} = y_{ij} - \varepsilon < y_{ij}$ to achieve cheating. In order to get a lower unreliability value below the threshold and make his/her cheating make sense in the service rating, user i needs to fabricate report y'_{ij} to make it close to $P_{m,j}$ and away from $1 - P_{f,j}$. With conditions $P_{f,i} < 0.5$ and $P_{m,i} < 0.5$, the smaller y'_{ij} is, the lower unreliability value will be get. On the other hand, the majority honest users trend to report the implicit opinion reports as $x_j = 1$ when $Q = h$. According to (9a) and (10a), the score of user i decreases with reducing y'_{ij} when his/her peer user j gives an accurate and honest report. Symmetrically, the dilemma still exists when $Q = l$. Therefore, it is difficult for malicious users to get high trustworthiness and low unreliability at the same time, if they report trickly. However, for those "bad functioning" honest users with relatively high error rate of judgement, the best choice is still reporting y_{ij} and y'_{ij} truthfully. It is unnecessary for them to modify their y'_{ij} because their benefit is the score and trustworthiness determined by the information and prediction reports, and this benefit is irrelevant to that whether their reports are accepted by the DPC or not.

B. Peer Prediction Based Service Rating

According to the user's trustworthiness and unreliability analysis above, we design the private-prior peer prediction based service rating method as following procedures.

1) For every user i who accepts the service, choose another non-overlapped user j randomly among his/her friends as i's peer.
2) Ask user i for his/her prior belief report $y_{ij} \in [0, 1]$, i.e. his/her peer j will provide a report to the cloud that j evaluates the service as high-quality.

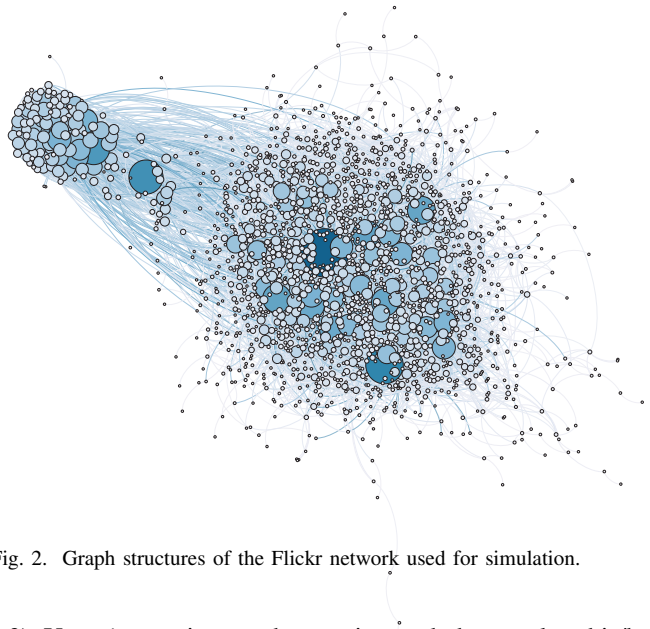


Fig. 2. Graph structures of the Flickr network used for simulation.

3) User i experiences the service and then makes his/her judgement $S_i = s_i$ for the quality of the service.
4) Ask user i for his/her posterior belief report $y'_{ij} \in [0, 1]$ to the cloud, with $y'_{ij} \neq y_{ij}$, that his/her peer j will provide a report of receiving a high-quality service.
5) The DPC calculates the unreliability of every user by applying (23), and removes reports of users with $\rho_i > \rho_{thr}$ from the service rating system.
6) The DPC infers the implicit opinion report x_i of user i through (8), and calculates user i's trustworthiness according to (9)/(10) and (11) assisted by user j's inferred opinion report x_j. Then remove reports of users with lowest trustworthiness from the service rating system.
7) The DPC makes the rating for the service by implicit opinion reports of users with both high trustworthiness and lower unreliability through (3).

V. Simulation Results

In this part, we perform numerical simulation experiments to analyze the properties and performances of the private-prior peer prediction service rating system and its influential factors such as the proportion of the malicious users and ε. First, we analyze the effect of the time accumulation on the trustworthiness and unreliability. In the peer prediction mechanism, we can notice that if the peer user of an honest user is a malicious one who decides to cheat when he/she reports to the cloud, the trustworthiness of the honest user trends to be low because of the strictly proper scoring rule. However, when malicious users are not predominant in the social network, which means that the proportion of the malicious users is less than half of the total, then honest users' accumulative trustworthiness will increase distinctly comparing with malicious ones in a long term.

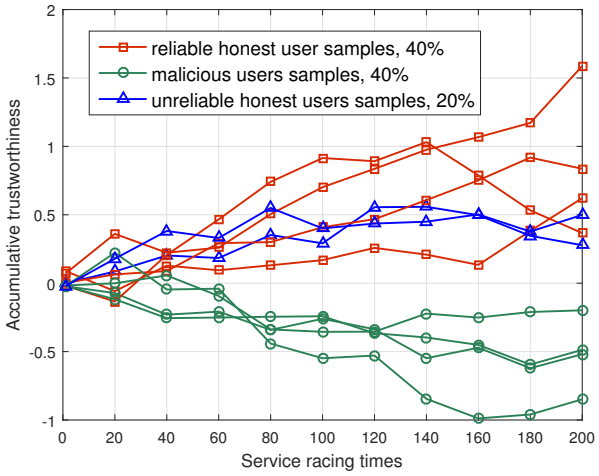
A. Simulation Settings

The simulation for the service rating system is operated based on the topology of Flickr, a real-world online social network database. The Flickr topology contains 5,899,882 edges connecting 80,513 users, and the edge represents the friendship of the connected two users. In addition, this friendship of users in the Flickr network, also known as the topology, is determined by their favorites. In other words, the connection between any two users is established if the corresponding two users are sharing the common favorites and have followed the same community. Then such two users will be considered as a pair of peers. The topology of the Flickr network are depicted in Fig. 2. These users are separated into three types, i.e., reliable honest users with high judgement accuracy rate, malicious users with high judgement accuracy rate and high error report rate, and unreliable honest users with relatively high judgement error rate but always report truthfully. The three types of users exist with some certain percentage. We set that false alarm of judgement P_{fa} and missed detection of judgement P_{md} are uniform distribution variables, and for all reliable and malicious users $P_{fa},\ P_{ma} \sim U\,[0.01, 0.02]$, and for unreliable users $P_{fa},\ P_{ma} \sim U\,[0.05, 0.06]$. In addition, as analyzed in the Remarks of Lemma 1, malicious users need to make sure that their false alarm and missed detection of report, P_f and P_m, are both smaller than 0.5 to achieve a continuous trick. Therefore, we set $P_f = P_m = 0.3$ (< 0.5) for malicious users in the following experiences. We assume that all honest users always report truthfully, i.e., $P_f = P_m = 0$.

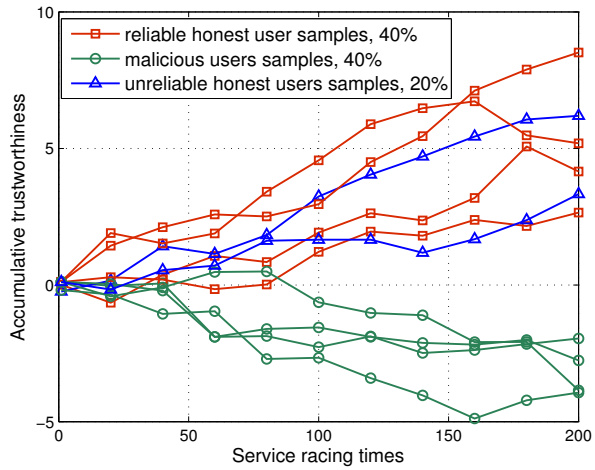
- *Historical database*. To calculate the unreliability of each user, the DPC needs to obtain their historic error rates of report. So we first establish the report database by allowing each user judge the quality of the service independently and then report to the cloud all according to the type of the user. The process repeats 80 times and in each time, the probability of high service quality is set as $P\,(Q = H) = 0.6$. In addition, the quality of the service is determined through the majority rule shown as (3) by applying reports from all of the users.

B. Accumulative Trustworthiness And Unreliability

Then in the following experiences, the private-prior peer prediction method is introduced, and the peer of each user is updated in every new experience. Then new implicit opinion reports (inferred by y_{ij} and y'_{ij}) and service rating results are added into the database and provide the historical data for the DPC. We consider that the trustworthiness and the unreliability of each user can be accumulated with the increasing service times. To calculate the trustworthiness, both of the scoring rules, i.e., binary logarithmic and binary quadratic, are applied. Simulation results of users' accumulative trustworthiness and unreliability in the following 200 times of service are shown in Fig. 3 and Fig. 4, in which the percentages of reliable honest user, malicious user and unreliable honest user are set as 40%, 40% and 20%, respectively. In both of the figures, we show the results of some sample users selected from the three types randomly. In Fig. 3, the trustworthiness of honest users might be negative at the beginning, when their peer users are the malicious. On the other hand, some malicious ones even obtain larger trustworthiness at the beginning, when their peers are also the malicious. However, resulting from the peer updating after each time of service, as well as the small proportion



(a) Binary logarithmic scoring



(b) Binary quadratic scoring

Fig. 3. The accumulative trustworthiness of user samples of three types.

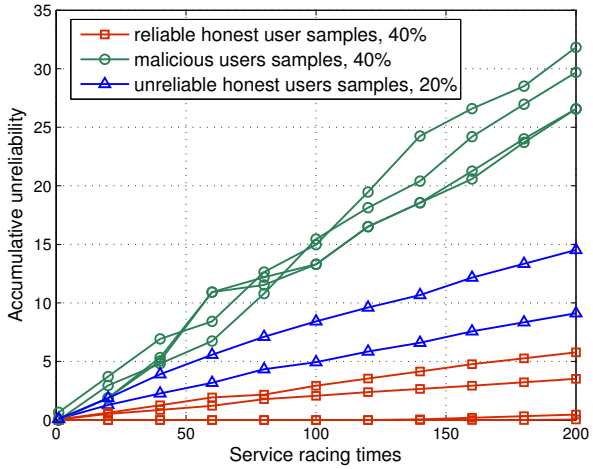


Fig. 4. The accumulative unreliability of user samples of three types.

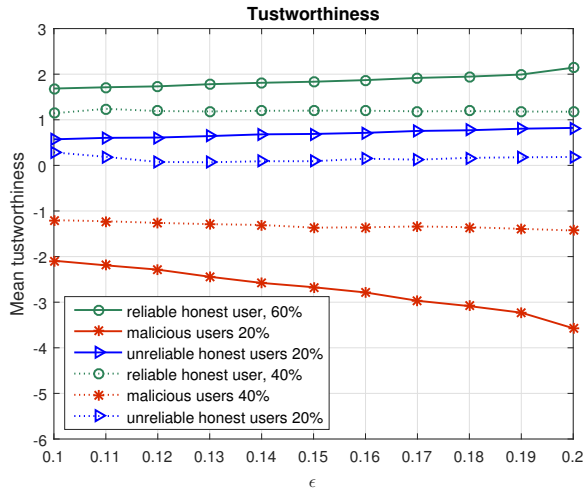
of the malicious, the predomination of honest users trends to work in a long term. Fig. 3 indicates that the accumulative trustworthiness of honest users grows with the service rating times or experience time. On the contrary, the accumulative trustworthiness of malicious users drops down and is negative. In addition, we can notice that no matter which scoring rule is applied, the accumulative trustworthiness shows the similar characteristics and tendency.

Similar results of accumulative unreliability are shown in Fig. 4, in which the gaps are more obvious among different types of users. Moreover, we can notice that unreliable honest users can be identified through the unreliability index, which cannot be achieved by the trustworthiness. This result demonstrates that the best choice for unreliable honest users is still reporting truthfully, and their unreliability will bring no hazard to their high positive trustworthiness.

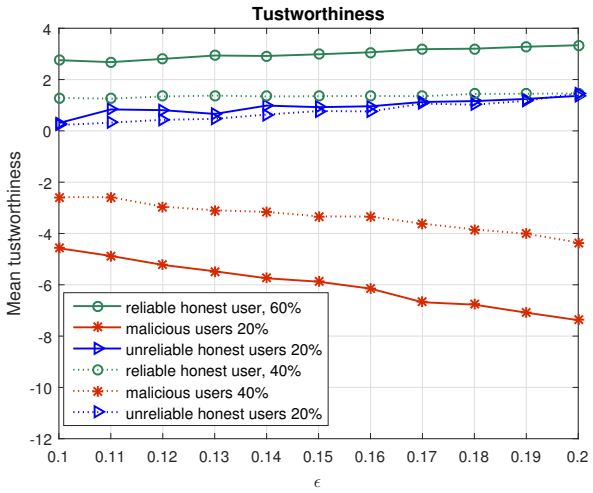
C. Influence of ε, Scoring Rules And User Structure

In the basic private-prior peer prediction mechanism, the strictly proper scoring rule leads malicious users to fabricate minimum ε, i.e., $y'_{ij} = y_{ij} + \varepsilon$ ($\varepsilon > 0$) when the quality of the service is low, and $y'_{ij} = y_{ij} - \varepsilon$ when the quality is high. In the trustworthy service racing system, the unreliability index proposed brings the dilemma to malicious users when they set ε as discussed previously. Next, we test the influence of ε on the average trustworthiness and unreliability. Considering two cases of user structure, the percentages of reliable honest user, malicious user and unreliable honest user are set as 60%, 20% and 20% in one case, respectively, and in another case are set as 40%, 40% and 20%. We repeat the service rating experiments for 200 times, and then calculate the average trustworthiness and unreliability of each type of users in these 200 times experiments (not the accumulative trustworthiness or unreliability). Results in Fig. 5(a) and Fig. 5(b) present the average trustworthiness when applying binary logarithmic and binary quadratic scoring rules, respectively, when $\varepsilon \in [0.1, 0.2]$. In addition, Fig. 6 presents how the average unreliability changes when ε increases. As depicted in Fig. 5 and Fig. 6, both the trustworthiness and unreliability decrease with the increase of ε for malicious users, which demonstrates the incentive and identification capabilities when combining trustworthiness and unreliability together to evaluate users reports. On the other hand, the average trustworthiness and unreliability of honest users are not sensitive to changing ε. In addition, we can notice that, when the percentage of malicious users is small, the gaps between the trustworthiness and unreliability malicious and honest user tend to be wide, which will make it much easier to identify the malicious.

Removing unreliable reports and reports from users with low trustworthiness, we rate the service quality by trustful reports to improve the accuracy of rating. In this part, we define the service rating accuracy as the ratio of the number of selected correct reports to the number of all correct reports. In addition, the threshold of unreliability is set as an empirical value obtained from the training of historical database, to be specific, $\rho_{thr} = 5$. Then we test the service rating accuracy over the proportion of malicious users, unreliable honest users' error rates of judgement and ε. Results shown in Fig. 7 indicate that the service rating accuracy decreases with the increasing proportion of malicious users. When this proportion



(a) Binary logarithmic scoring



(b) Binary quadratic scoring

Fig. 5. The average trustworthiness of different types of users versus the percentage of each user type and ε.

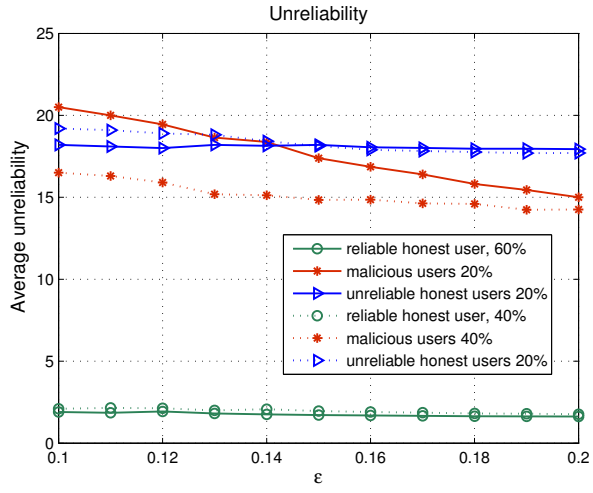


Fig. 6. The average unreliability of different types of users versus the percentage of each user type and ε.

is closed to 0.5, the rating accuracy decreases distinctly because of the probable cooperative cheating. In addition,

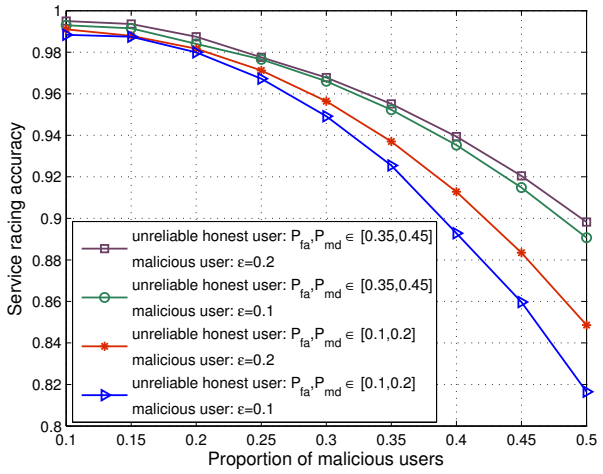


Fig. 7. The service racing accuracy versus the the percentage of each user type, error rates of judgement P_{fa}, P_{md} and ε.

the rating accuracy is higher when unreliable honest users' $P_{fa},\ P_{ma} \sim U[0.35, 0.45]$ than $P_{fa},\ P_{ma} \sim U[0.1, 0.2]$, which results from that honest users with higher judgement error rates can be identified more easily by applying the unreliability index. Fig. 7 also indicates that the lower ε malicious users set, the harder they can be detected through the trustworthiness.

VI. Conclusion

In this paper, we proposed an cloud based architecture for the service rating system. To achieve a trustworthy service rating, a private-prior peer prediction based mechanism was designed to identify malicious and dishonest users. Coupled with some certain strictly proper scoring rules, the peer prediction method can evaluate users' trustworthiness and motivate them to report honestly. Moreover, an unreliability index was also designed to ensure the reliability of the users' reports. According to the trustworthiness and unreliability index, untruthful and unreliable reports can be identified and eliminated to improve the accuracy of service rating. Simulation results indicated that the proposed peer prediction based trustworthy service rating system can identify malicious and unreliable behaviours effectively, and achieve relatively high service rating accuracy.

Appendix A
Proof of Lemma 1

Proof: In (4),

$$\begin{aligned} & y'_{ij}(h) = P(x_j = 1 \,|S_i = h) \\ &= P(x_j = 1 \,|Q = h)\, P(Q = h \,|S_i = h) \\ &\quad + P(x_j = 1 \,|Q = l)\, P(Q = l \,|S_i = h) \\ &= \frac{1}{P(S_i = h)} \left[(1 - P_{f,j})(1 - P_{fa,i}) P(Q = h) \right. \\ &\quad \left. + P_{m,j} P_{md,i} P(Q = l)\right]. \end{aligned} \tag{24}$$

Similarly,

$$\begin{aligned} & y'_{ij}(l) = P(x_j = 1 \,|S_i = l) \\ &= P(x_j = 1 \,|Q = h)\, P(Q = h \,|S_i = l) \\ &\quad + P(x_j = 1 \,|Q = l)\, P(Q = l \,|S_i = l) \\ &= \frac{1}{P(S_i = l)} \left[(1 - P_{f,j})\, P_{fa,i} P(Q = h)\right. \\ &\quad \left. + P_{m,j}(1 - P_{md,i})\, P(Q = l)\right]. \end{aligned} \tag{25}$$

So y_{ij} can be written as

$$y_{ij} = (1 - P_{f,j})\, P(Q = h) + P_{m,j} P(Q = l). \tag{26}$$

Then,

$$\begin{aligned} & y'_{ij}(h) - y_{ij} \\ &= \frac{1}{P(S_i = h)} \left[(1 - P_{f,j})(1 - P_{fa,i})\, P(Q = h)\right. \\ &\quad \left. + P_{m,j} P_{md,i} P(Q = l)\right] - (1 - P_{f,j})\, P(Q = h) \\ &\quad - P_{m,j} P(Q = l) \\ &= \frac{1}{P(S_i = h)} \{(1 - P_{f,j})(1 - P_{fa,i})\, P(Q = h) \\ &\quad + P_{m,j} P_{md,i} P(Q = l) \\ &\quad - [(1 - P_{f,j})\, P(Q = h) + P_{m,j} P(Q = l)] \\ &\quad \cdot [(1 - P_{fa,i})\, P(Q = h) + P_{md,i} P(Q = l)]\} \\ &= A_0(Q, S_i) \left[(1 - P_{f,j})(1 - P_{fa,i}) + P_{m,j} P_{md,i}\right. \\ &\quad \left. - P_{m,j}(1 - P_{fa,i}) - (1 - P_{f,j})\, P_{md,i}\right] \\ &= A_0(Q, S_i)(1 - P_{fa,i} - P_{md,i})(1 - P_{f,j} - P_{m,j}), \end{aligned} \tag{27}$$

where

$$A_0(Q, S_i) = \frac{P(Q = h)\, P(Q = l)}{P(S_i = h)}. \tag{28}$$

Therefore, when $P_{fa,i} + P_{md,i} < 1$ and $P_{f,j} + P_{m,j} < 1$, inequality $y'_{ij}(h) > y_{ij}$ holds. By symmetry, we have $y'_{ij}(l) < y_{ij}$ under the same conditions.

This completes the proof of Lemma 1.

REFERENCES

[1] M. Chen, Y. Zhang, L. Hu, T. Taleb, and Z. Sheng, "Cloud-based wireless network: Virtualized, reconfigurable, smart wireless network to enable 5G technologies," *Mobile Networks and Applicat.*, vol. 20, no. 6, pp. 704–712, Feb. 2015.

[2] G. Zhao, X. Qian, and C. Kang, "Service rating prediction by exploring social mobile users' geographical locations," *IEEE Trans. Big Data*, vol. 3, no. 1, pp. 67–78, Mar. 2017.

[3] J. Du, C. Jiang, S. Yu, K. C. Chen, and Y. Ren, "Privacy protection: A community-structured evolutionary game approach," in *IEEE Global Conf. Signal Inform. Process. (GlobalSIP)*. Washington, DC, USA, 7-9 Dec. 2016, pp. 415–419.

[4] G. Zhao, X. Qian, and X. Xie, "User-service rating prediction by exploring social users' rating behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 496–506, Mar. 2016.

[5] X. Li, H. Ma, F. Zhou, and W. Yao, "T-broker: A trust-aware service brokering scheme for multiple cloud collaborative services," *IEEE Trans. Inf. Forens. Security*, vol. 10, no. 7, pp. 1402–1415, Jul. 2015.

[6] Y. Yang, Y. Sun, S. Kay, and Q. Yang, "Securing rating aggregation systems using statistical detectors and trust," *IEEE Trans. Inf. Forens. Security*, vol. 4, no. 4, pp. 883–898, Dec. 2009.

[7] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu, "Privacy or utility in data collection? a contract theoretic approach," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1256–1269, Oct. 2015.

[8] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE Trans. Broadcasting*, vol. 57, no. 2, pp. 165–182, Jun. 2011.

[9] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. Image Processing*, vol. 19, no. 6, p. 1427, 2010.

[10] X. Liang, X. Lin, and X. S. Shen, "Enabling trustworthy service evaluation in service-oriented mobile social networks," *IEEE Trans. Parallel Distrib.uted Syst.*, vol. 25, no. 2, pp. 310–320, Feb. 2014.

[11] B. Jennings and P. Malone, "Flexible charging for multi-provider composed services using a federated, two-phase rating process," in *2006 10th IEEE/IFIP Network Operations Manage. Symp. (NOMS 2006)*. Vancouver, BC, Apr. 2006, pp. 13–23.

[12] S. Sharma and J. Negi, "Customer importance rating of service quality dimensions for automobile service," *Int. J. of Eng. Sci. and Technology*, vol. 6, no. 12, pp. 822–826, Dec. 2014.

[13] G. Di Fabbrizio, A. Aker, and R. Gaizauskas, "Summarizing online reviews using aspect rating distributions and language modeling," *IEEE Intell. Syst.*, no. 3, pp. 28–37, May. 2013.

[14] M. Allahbakhsh and A. Ignjatovic, "An iterative method for calculating robust rating scores," *IEEE Trans. Parallel Distrib.uted Syst.*, vol. 26, no. 2, pp. 340–350, Feb. 2015.

[15] K. Nagaraja, G. Gama, R. Bianchini, R. P. Martin, W. Meira Jr, and T. D. Nguyen, "Quantifying the performability of cluster-based services," *IEEE Trans. Parallel Distrib.uted Syst.*, vol. 16, no. 5, pp. 456–467, May 2005.

[16] X. Ye and J. Zheng, "An adaptive rating system for service computing," in *2013 12th IEEE Int. Conf. Trust, Security, Privacy Computing Commun. (TrustCom)*. Melbourne, VIC, Jul. 2013, pp. 1817–1824.

[17] G. Zhao, X. Qian, and X. Xie, "User-service rating prediction by exploring social users' rating behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 496–506, Mar. 2016.

[18] J. Du, C. Jiang, E. Gelenbe, Z. Han, Y. Ren, and M. Guizani, "Networked data transaction in mobile networks: A prediction-based approach using auction," in *IEEE Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*. Limassol, Cyprus, 25-29 Jun. 2018.

[19] Y. Kim, E.-W. Jhee, J. Choe, J.-S. Choi, and Y. Shin, "A measurement model for trustworthiness of information on social network services," in *2015 Int. Conf.Inform. Networking (ICOIN)*. Cambodia, Jan. 2015, pp. 437–438.

[20] J. Du, C. Jiang, K. C. Chen, Y. Ren, and H. V. Poor, "Community-structured evolutionary game for privacy protection in social networks," *IEEE Trans. Inf. Forens. Security*, vol. 13, no. 3, pp. 574–589, Mar. 2018.

[21] J. Jiang, G. Han, F. Wang, L. Shu, and M. Guizani, "An efficient distributed trust model for wireless sensor networks," *IEEE Trans. Parallel Distrib.uted Syst.*, vol. 26, no. 5, pp. 1228–1237, May 2015.

[22] D. He, S. Chan, and M. Guizani, "User privacy and data trustworthiness in mobile crowd sensing," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 28–34, Feb. 2015.

[23] B. Kantarci and H. T. Mouftah, "Trustworthy sensing for public safety in cloud-centric internet of things," *IEEE J. Internet of Things*, vol. 1, no. 4, pp. 360–368, Jul. 2014.

[24] N. Zhang, N. Cheng, N. Lu, H. Zhou, J. W. Mark, and X. Shen, "Risk-aware cooperative spectrum access for multi-channel cognitive radio networks," *IEEE J. Sel. Areas. Commun.*, vol. 32, no. 3, pp. 516–527, Mar. 2014.

[25] T. Noor, Q. Sheng, L. Yao, S. Dustdar, and A. Ngu, "CloudArmor: Supporting reputation-based trust management for cloud services," *IEEE Trans. Parallel Distrib.uted Syst.*, vol. 27, no. 2, pp. 367–380, Mar. 2015.

[26] B. Li, L. Liao, H. Leung, and R. Song, "PHAT: A preference and honesty aware trust model for web services," *IEEE Trans. Network and Service Manage.*, vol. 11, no. 3, pp. 363–375, Sept. 2014.

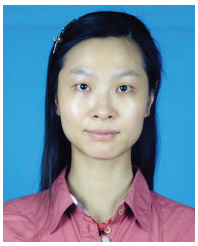
[27] H. Dong, C. Wu, Z. Wei, and Y. Guo, "Dropping activation outputs with localized first-layer deep network for enhancing user privacy and data security," *IEEE Trans. Inf. Forens. Security*, vol. 13, no. 3, pp. 662–670, Mar. 2018.

[28] L. Xu, C. Jiang, N. He, Z. Han, and A. Benslimane, "Trust-based collaborative privacy management in online social networks," *IEEE Trans. Inf. Forens. Security*, pp. 1–1, May 2018.

[29] B. Rashidi, C. Fung, A. Nguyen, T. Vu, and E. Bertino, "Android user privacy preserving through crowdsourcing," *IEEE Trans. Inf. Forens. Security*, vol. 13, no. 3, pp. 773–787, Mar. 2018.

[30] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu, "User participation in collaborative filtering-based recommendation systems: A game theoretic

approach," *IEEE Trans. Cybernetics*, vol. PP, no. 99, pp. 1–14, Feb. 2018.
[31] J. Du, C. Jiang, Z. Han, H. Zhang, S. Mumtaz, and Y. Ren, "Contract mechanism and performance analysis for data transaction in mobile social networks," *IEEE Trans. Network Sci. & Eng.*, vol. PP, no. 99, pp. 1–1, Dec. 2017.
[32] R. Selten, "Axiomatic characterization of the quadratic scoring rule," *Experimental Econ.*, vol. 1, no. 1, pp. 43–62, Jun. 1998.
[33] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, Jan. 2012.
[34] N. Miller, P. Resnick, and R. Zeckhauser, "Eliciting informative feedback: The peer-prediction method," *Manage. Sci.*, vol. 51, no. 9, pp. 1359–1373, Sept. 2005.
[35] J. Witkowski and D. C. Parkes, "Peer prediction without a common prior," in *Proc. 13th ACM Conf. Electron. Commerce*. ACM, Valencia, Spain, Jun. 2012, pp. 964–981.
[36] B. Faltings, J. J. Li, and R. Jurca, "Incentive mechanisms for community sensing," *IEEE Trans. Comput.*, vol. 63, no. 1, pp. 115–128, Jul. 2014.
[37] G. Radanovic and B. Faltings, "Incentives for truthful information elicitation of continuous signals," in *Twenty-Eighth AAAI Conf. Artificial Intell.*, Jul. 2014, pp. 770–776.
[38] Y. Gan, C. Jiang, N. C. Beaulieu, J. Wang, and Y. Ren, "Secure collaborative spectrum sensing: A peer-prediction method," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4283–4294, Oct. 2016.
[39] J. Du, E. Gelenbe, C. Jiang, H. Zhang, and Y. Ren, "Contract design for traffic offloading and resource allocation in heterogeneous ultra-dense networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2457–2467, Nov. 2017.
[40] X. A. Gao, A. Mao, Y. Chen, and R. P. Adams, "Trick or treat: putting peer prediction to the test," in *Proc. fifteenth ACM conf. Econ. computation*. California, US, Jun. 2014, pp. 507–524.
[41] J. Du, C. Jiang, J. Wang, S. Yu, and Y. Ren, "Trustable service rating in social networks: A peer prediction method," in *IEEE Global Conf. Signal Inform. Process. (GlobalSIP)*. Washington, DC, USA, 7-9 Dec. 2016, pp. 415–419.
[42] J. Witkowski, Y. Bachrach, P. Key, and D. C. Parkes, "Dwelling on the negative: Incentivizing effort in peer prediction," in *First AAAI Conf. Human Computation Crowdsourcing*, Nov. 2013, pp. 190–197.
[43] G. Radanovic and B. Faltings, "A robust bayesian truth serum for non-binary signals," in *Proc. 27th AAAI Conf. Artificial Intell. (AAAI 2013)*, no. EPFL-CONF-197486, Jul. 2013, pp. 833–839.
[44] J. Du, E. Gelenbe, C. Jiang, H. Zhang, Z. Han, and Y. Ren, "Data transaction modeling in mobile networks: Contract mechanism and performance analysis," in *IEEE Global Commun. Conf. (GLOBECOM)*. Singapore, 4-8 Dec. 2017.
[45] J. Du, E. Gelenbe, C. Jiang, Z. Han, Y. Ren, and M. Guizani, "Cognitive data allocation for auction-based data transaction in mobile networks," in *IEEE Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*. Limassol, Cyprus, 25-29 Jun. 2018.



Jun Du (S'16-M'18) received the B.S. degree in information and communication engineering from the Beijing Institute of Technology in 2009, and the M.S. and Ph.D. degrees in information and communication engineering from Tsinghua University, Beijing, in 2014 and 2018, respectively. From 2016 to 2017, she was a Sponsored Researcher and visited the Imperial College London. She currently holds a post-doctoral position with the Department of Electrical Engineering, Tsinghua University. Her research interests are mainly in resource allocation and system security of heterogeneous networks and space-based information networks.

Erol Gelenbe (S'67-M'70-SM'79-F'86-LF'11) is a Fellow of ACM. He is the Dennis Gabor Professor in Electrical and Electronic Engineering at Imperial College, London. He received his BS from the Middle East Technical University, Ankara, Turkey, his MS (1968) and PhD (1970) from the Polytechnic Institute of Brooklyn (now New York University) under Professor Edward J. Smith Jr., and received the Doctorat d'État ès Sciences Mathématiques from Université Pierre et Marie Curie, Paris, under Prof. Jacques Louis-Lions. Also awarded Honorary Doctorates from the Università di Roma II, Italy (1996), Boğaziçi University, Istanbul, Turkey (2004) and Université de Liège, Belgium (2006), he received the Parlar Foundation Science Award, Turkey (1994), Grand Prix France Télécom of the French Academy of Sciences (1996), the ACM-SIGMETRICS Life-Time Achievement Award (2008), IET Oliver Lodge Medal (2010), the "In Memoriam Dennis Gabor Prize" from the Hungarian Academy of Sciences (2013), and the Mustafa Prize of the Pardis Foundation (2017). He was elected a Fellow of Academia Europaea (2005), of the French National Academy of Technologies (2008), and the Science Academy of Turkey (2012).

Chunxiao Jiang (S'09-M'13-SM'15)received the B.S. degree (Hons.) in information engineering from Beihang University in 2008 and the Ph.D. degree (Hons.) in electronic engineering from Tsinghua University in 2013. From 2013 to 2016, he held a post-doctoral position with the Department of Electronic Engineering, Tsinghua University, during which he visited the University of Maryland College Park, College Park, MD, USA, and The University of Southampton. He is currently an Assistant Professor with the Tsinghua Space Center, Tsinghua University. He was a recipient of the IEEE Globecom Best Paper Award in 2013, the IEEE GlobalSIP Best Student Paper Award in 2015, and the IEEE Communications Society Young Author Best Paper Award in 2017.

Haijun Zhang (M'13-SM'17) is currently a Full Professor in University of Science and Technology Beijing, China. He was a Postdoctoral Research Fellow in Department of Electrical and Computer Engineering, the University of British Columbia (UBC), Vancouver Campus, Canada. He serves as Editor of IEEE Transactions on Communications, IEEE Transactions on Green Communications and Networking, IEEE 5G Tech Focus. He received the IEEE CSIM Technical Committee Best Journal Paper Award in 2018 and IEEE ComSoc Young Author Best Paper Award in 2017.

Yong Ren (SM'16) received his B.S, M.S and Ph.D. degrees in electronic engineering from Harbin Institute of Technology, China, in 1984, 1987, and 1994, respectively. He worked as a post doctor at Department of Electronics Engineering, Tsinghua University, China from 1995 to 1997. Now he is a professor of Department of Electronics Engineering and the director of the Complexity Engineered Systems Lab (CESL) in Tsinghua University. His current research interests include complex systems theory and its applications to the optimization and information sharing of the Internet, Internet of Things and ubiquitous network, cognitive networks and Cyber-Physical Systems.

H. Vincent Poor (S'72-M'77-SM'82-F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering. From 2006 until 2016, he served as Dean of Princetons School of Engineering and Applied Science. He has also held visiting appointments at several other institutions, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory and signal processing, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society and other national and international academies. He received the Technical Achievement and Society Awards of the IEEE Signal Processing Society in 2007 and 2011, respectively. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships from Peking University and Tsinghua University, both conferred in 2017, and a D.Sc. honoris causa from Syracuse University, awarded in 2017.