This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

Information Credibility Modeling in Cooperative Networks: Equilibrium and Mechanism Design

Chunxiao Jiang, *Senior Member, IEEE*, Linling Kuang, *Member, IEEE*, Zhu Han, *Fellow, IEEE*, Yong Ren, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—In a cooperative network the user equipment (UE) share information with each other for cooperatively achieving a common goal. However, owing to the concerns of privacy or cost, UEs may be reluctant to share genuine information, which raises the information credibility problem addressed. Diverse techniques have been proposed for enhancing the information credibility in various scenarios. However, there is paucity of information on modeling the UEs' decision making behavior, namely as to whether they are willing/able to share genuine information, even though this directly affects the information credibility across the network. Hence, we propose a game theoretic framework for the associated information credibility modelling by taking into account the users' information sharing strategies and utilities. This framework is investigated under both a homogeneous model and a heterogeneous model. The spontaneous information credibility equilibria of both models are derived and analyzed, including the closed-form analysis of the homogeneous model based on a sophisticated evolutionary game model and on the reinforcement learning based analysis of the heterogeneous model. Moreover, a credit mechanism is designed for encouraging the UEs to share genuine information. Experimental results relying on real-world data traces support our utility function formulation, while our simulation results verify the theoretical analysis and show that all UEs are encouraged by the proposed algorithm to share genuine information with a probability of one, when a credit mechanism is invoked. The proposed modelling techniques may be applied in diverse cooperative networks, including classic wireless networks, vehicular networks, as well as social networks.**

***Index Terms*—Information credibility, cooperative networks, game theory, reinforcement learning, self-organizing networks.**

I. Introduction

At the time of writing, cooperation is becoming a pervasive phenomenon in various networks, where the users or equipment sharing the same interest may embark on cooperation to achieve a common target [1]. In a distributed self-organizing network, each individual has limited access to the global network's status, but cooperation is capable of assisting them in enhancing their judgement and decision making. For instance, in cognitive radio networks, cooperative spectrum sensing can be invoked by a group of secondary users (SUs) for enhancing the detection probability of the primary users (PUs) [2]; in vehicular networks, vehicles may cooperatively share the location or traffic status by information dissemination in order to enhance both the security and the traffic flow [3]; in cellular networks, the base stations may cooperatively serve all user equipment (UE) by sharing their channel information, which improves the interference cancellation performance and increases the spatial multiplexing gain [4]; in a heterogeneous network, different network operators may cooperatively manage the UEs' access for the sake of efficient load balancing, for example by combining visible light communications in the downlink with WiFi in the uplink [5].

In support of cooperation, one of the most crucial issues is the information exchange/sharing, especially in distributed scenarios [6]. Since each network entity can only acquire local information, efficient information exchange enhances the cooperation among neighbors. However, information sharing is not gratuitous for each individual, in fact it may be quite expensive under some circumstances. In such a case, due to the natural selfishness and rational inclination of UEs in the network, they may be reluctant to share genuine information, for example for the sake of privacy preservation or cost saving. In this paper, we propose a generalized game-theoretic information credibility modelling framework for cooperative networks by studying the UEs' incentives and behaviors in information sharing. This framework is expected to reveal how a group of UEs adjust their individual actions during the network's operation, and how they can spontaneously achieve convergence to a stable equilibrium after a few rounds of interactions. The contribution of this paper can be summarized as follows.

1) A homogeneous information sharing model is analyzed based on a sophisticated evolutionary game, where the information sharing strategies of all UEs are identical. We provide the closed-form analysis of the evolutionarily stable equilibrium and quantify the specific proportion of UEs who would like to share genuine information.
2) Additionally, a more realistic heterogeneous model is also analyzed, where the information sharing strategies of all UEs are different. Based on the classic reinforcement learning model, we characterize the heterogenous UEs' spontaneous information credibility equilibrium.
3) Moreover, a credit mechanism is designed for encouraging the UEs to actively share genuine information so as to enhance the information credibility. Experiments

This work was partially supported by the National Nature Science Foundation of China (Grant Nos. 61371079, 91438206, 91338108, and 61231011), the National Basic Research Program of China (Grant No. 2013CB329001).

C. Jiang and L. Kuang (*corresponding author*) are with the Tsinghua Space Center, Tsinghua University, Beijing China (email:{jchx,kll}@tsinghua.edu.cn).

Z. Han is with Electrical and Computer Engineering Department as well as Computer Science Department, University of Houston, Houston, TX, USA (email: zhan2@uh.edu).

Y. Ren is with the Department of Electronic Engineering, Tsinghua University, Beijing, China (email: reny@tsinghua.edu.cn).

L. Hanzo is with the School of Electronic and Computer Science, University of Southampton, Southampton, SO17 1BJ U.K (email: lh@ecs.soton.ac.uk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

based on real-world data traces are conducted to verify our utility formulation and simulations are conducted to verify our theoretical analysis.

The rest of the paper is organized as follows. Our system model is introduced in Section II. Based on the system model, we present the homogeneous and heterogeneous information sharing models in Sections III and IV, respectively. Each of these sections commences with the utility function definition, and then we analyze the credibility equilibrium, as well as the credit mechanism conceived in Section IV. Section V provides our simulation results and discussions. Finally, we conclude the paper in Section VI.

II. Related Works

The information credibility issue has already been investigated in various application scenarios [12]-[23]. Specifically, in the early Internet era, peer-to-peer (P2P) networks encountered malicious recommendations and file sharing problems, as exemplified by P2P transactions in [7] and by P2P virus proliferation in [8]. Song *et al.* [7] proposed a P2P reputation system based on fuzzy logic aided inferences by aggregating peer reputations, which can better handle uncertainty, fuzziness and incomplete information in peer-trust reports. Similarly, the peer-trust reputation can also be used for limiting the P2P virus prorogation problem of file exchanges [8]. The authors of [9] proposed a credibility model by storing and evaluating each witness' past testimony-reporting, which effectively mitigated the adverse influence of unfair testimonies. The credibility was also a prominent issue in social swarming, where the smart-phones can be requested by a center to report on events observed in the physical world [10]. From the game theoretic perspective, Canzian *et al.* [11] studied a general interaction model between a system designer and a group of users processing private information that the designer does not have.

As for the qualitative analysis of credibility, Sun *et al.* [12] utilized information theoretic techniques for quantifying the information credibility and proposed a trusted routing protocol for *ad hoc* networks. An information credibility model conceived for wireless multicasting was studied in the context of wireless network scenarios in [13], while the security versus reliability analysis of opportunistic relaying invoked in cooperative communication networks was carried out in [14]. In cognitive radio networks, a protocol was devised for mitigating the primary user emulation attacks and for reducing the call-dropping probability [15]. Furthermore, the concept of physical layer security was proposed for reducing the information entropy leakage in wireless networks [16]. In the context of the emerging topics of e-Health and the Internet of Things (IoT), the information trustworthiness issues were studied in [17] and [18], where the methodology advocated exploited both the subjective recommendations and the objective observations gleaned.

The other important issue in cooperative networks is cooperation stimulation. In the literature, efforts have been made for mathematically analyzing cooperation using game theory. Chen *et al.* [19] proposed a general cooperation stimulation mechanism based on indirect reciprocity theory, while Niu *et al.* [20] proposed a cooperation stimulation strategy for wireless multi-cast networks based on infinite repeated game theory. To elaborate a little bit further, typically a credit mechanism is involved in the modeling of cooperative stimulation, such as the reputation based credit mechanism of cognitive networks [21], the credit-based reward scheme of delay-tolerant networks [22] and the credit pre-reservation mechanism of online charging systems [23]. Similar to the cooperation stimulation, the model employed in this paper can be regarded as a genuine information sharing stimulation. In contrast to the existing credit based cooperation stimulation mechanisms, we associate the interaction probability with the credit in this paper, which can restrict the interaction frequency of the users having a low reputation, who would have limited access to the others' information.

Generally, all the aforementioned contributions were focused on designing specific algorithms for particular systems in the interest of enhancing the information credibility or trustworthiness attained. However, the role of the UEs' actions and decision making in dynamic information sharing have not been taken into account, even though the grade of information credibility is directly determined by the UEs' decisions on whether to share genuine information with each other. This dilemma may be resolved with the aid of game theory by modelling the learning and decision making problem by relying on each UE's actions and on the utility of information sharing in the network. Moreover, the existing contributions on information credibility [12]-[22] approached the problem from a bottom-up perspective, i.e. by considering a particular aspect in a specific scenario, whilst there is no general framework that models this problem in a top-down manner. Against this background, in this paper, we propose a generalized game-theoretic information credibility modelling framework for cooperative networks by studying the UEs' incentives and behaviors in information sharing.

III. System Model

We consider a cooperative network supporting N selfish UEs numbered $\{1, 2, ..., N\}$, each aiming for maximizing its own utility. The UEs have the capability of acquiring new information and are willing to share their information with each other in order to make better-informed decisions. In this model, it is assumed that the more trusted the information a UE relies on, the better its decisions become. As shown in Fig. 1, at the beginning of a time period, each UE acquires new information with a probability of p_a, where p_a is a decreasing function of the information acquisition cost. Given that new information has indeed been obtained, the UE has to decide whether it will truthfully and authentically share this information with others. Alternatively, whether it will manipulate the shared information to render it useless either due to privacy concerns or with the objective of gaining an unfair resource allocation advantage. Therefore, although all the UEs act in a cooperative manner in the network, they occasionally may share random or manipulated information for the sake of improving their own utility. Following the information sharing phase, each UE

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

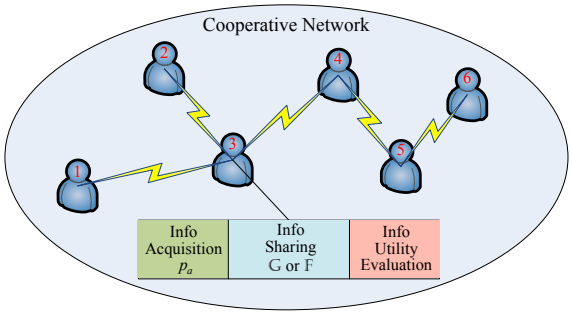

Fig. 1. System model of a cooperative network.

utilizes its own information, as well as the shared information to make a decision independently. Then, at the end of the time period, the UE evaluates the performance attained as a result of its decision and then adjusts its actions in preparation for the next round. Here, we consider a practical scenario where a UE is unable to ascertain the credibility of the shared information gleaned, until the information is actually used for its decision making and the resultant performance is actually evaluated.

Based on the aforementioned model, each UE can be considered to have two possible actions/strategies, namely that of sharing genuine or false information, which can be denoted by $\mathbb{G}$ and $\mathbb{F}$, respectively. It is possible that even if a UE shared its genuine information, the quality of the information is unsatisfactory and hence leads to a low utility. However, in this paper, we do not consider this information quality issue, rather we focus our attention on the information credibility from a security perspective. As a matter of fact, the above-mentioned low-quality information scenario can also be regarded as adopting strategy $\mathbb{F}$ in our model. The mixed strategy of each UE can be defined by $P_{i,\mathbb{G}}$, representing the probability of UE i opting for sharing genuine information, and with $1 - P_{i,\mathbb{G}}$ being the probability of supplying false or useless information. Since a UE utilizes all the shared information to make a decision, the amount of utility gleaned by each UE depends on all other UEs' strategies as to whether it shares genuine information. Moreover, as mentioned, sharing genuine information incurs some additional costs, which is denoted by c_i for UE i.

In the following sections, we will consider two models: a homogeneous and a heterogenous model defined as follows.

- *Homogeneous model:* Every UE relies on the same mixed strategy, but the specific manifestation of everyone's pure strategy in each time period can be different from each other, as quantified by $P_{i,\mathbb{G}}$. Meanwhile, each UE adopts the same strategy in its interaction with all other UEs without any discrimination, i.e. either sharing genuine information with all others or sharing false information with all others.
- *Heterogeneous model:* All UEs have different mixed strategies and each UE adopts different strategies in its interaction with different UEs. Hence UE i may share genuine information with UE j, but false information with UE k.

Again, we will conceive a general information sharing model, without concentrating on the specific form of the information and its utility. For instance, in cognitive radio networks, the information can be represented by the energy-detection samples, while the utility can be constituted by the associated detection probability; in vehicular networks, the information can be a specific location and the utility can be the corresponding traffic status; in social networks, the information may be constituted by recommendations of a particular commodity, while the utility can be the resultant user experience.

IV. Homogeneous Model: Spontaneous Credibility Equilibrium

In this section, we analyze the homogeneous information sharing model in a cooperative network, where all the UEs are identical in terms of both their mixed strategy and their information acquisition/sharing costs and rewards. Hence, we can omit the UE index in this section, by simply denoting the mixed strategy as $P_{\mathbb{G}}$ and information acquisition cost as c. Note that when the number of UEs is sufficiently high, a specific UE's probability of sharing genuine information, $P_{\mathbb{G}}$, is equivalent to the percentage of UEs adopting strategy $\mathbb{G}$ in the network. As we will demonstrate, this homogeneous model may assist us in deriving a closed-form analysis of both the credibility evolution, as well as of the steady-state percentage of UEs sharing genuine information. In the following, we will first define the utility function of information sharing, and then analyze the evolutionary equilibrium of the dynamic information sharing process.

A. Utility Formulation

In a cooperative network, each UE utilizes the shared information gleaned from each other to make decisions, where the utility is closely related to the other UEs' information sharing actions, i.e. whether they are willing to share genuine information. Assuming that there are n UEs sharing genuine information at a specific time instant, the reward that each UE can obtain by utilizing the information should be proportional to n due to assuming that the positive externality property holds, which we define as a non-decreasing function $f(n)$ of n. Naturally, in different application scenarios, the form of $f(n)$ is different. Again, for instance, in cooperative spectrum sensing of cognitive radio networks, $f(n)$ should be the probability of SUs' correctly detecting the PUs' signals, which can be formulated as [24]

$$f(n) = Q\left(-\left(1+\gamma_s-\frac{E}{\sigma_n^2}\right)\sqrt{\frac{n\lambda_s T_s}{2\gamma_s+1}}\right), \qquad (1)$$

where E is the threshold of the energy detector, σ_n^2 is the variance of the additive white Gaussian noise, γ_s is the signal-to-noise ratio given by the PU-power to SU-noise ratio when the PUs are present, λ_s is the SUs' sampling rate and T_s is the sampling period, $Q(x)$ is the Q-function. From (1), we can see that the more users share genuine spectrum sensing information, the higher the detection probability becomes. Since the probability is bounded by $[0, 1]$, diminishing returns would appear, when the number of users sharing genuine information is sufficiently large. For Doppler-based cooperative positioning

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

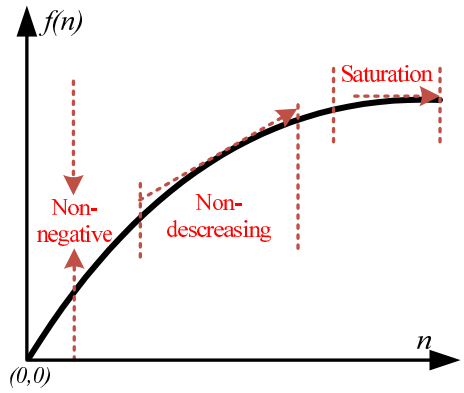

Fig. 2. Characteristic of the utility function.

in vehicle networks, $f(n)$ should be the positioning accuracy, which can be evaluated by the Fisher information as [25]

$$f(n) = \frac{1}{\sigma_k^2} + \frac{1}{\sigma_\omega^2} \sum_{i=1}^{n} \left[\frac{\partial \omega_i}{\partial k} \right]^2, \tag{2}$$

where k may represent either the x-direction or y-direction, σ_k^2 is the variance of the GPS-based location observation noise, ω_i is the Doppler shift of the signal with respect to its neighbour i, and σ_ω^2 is the variance of the Doppler shift observation error, which is assumed to be Gaussian. From (2), we can see that the more users share genuine location information, the higher the positioning accuracy becomes. Similarly, since the accuracy is bounded by $[0, 1]$, diminishing returns are achieve when the number of users sharing genuine position information is high. For collaborative filtering in an item recommender system in social networks, $f(n)$ should be the prediction accuracy, which can be quantified by the mean absolute prediction error as [26]

$$f(n) = \frac{1}{|\mathcal{J}|} \sum_{j \in \mathcal{J}} \left| \bar{r}_u + \frac{\sum_{i=1}^{n} s_{u,i}(r_{i,j} - \bar{r}_i)}{\sum_{i=1}^{n} |s_{u,i}|} - r_{u,j} \right|, \tag{3}$$

where $\mathcal{J}$ is the item set, $\bar{r}_u$ is the UE's overall average rating, $s_{u,i}$ is the similarity between a user and his/her neighbour i in the network, $r_{i,j}$ is UE i's rating regarding item j, whilst finally, $r_{u,j}$ is the user's true rating in the testing dataset, i.e. outside the training set. From (3), we can see that the more users share genuine recommendations, the lower the prediction error becomes, which is also bounded by $[0, 1]$.

Again, although the specific forms of $f(n)$ can be different in diverse scenarios, it should obey the following properties in most scenarios, as augmented by Fig. 2,

1) non-negative property, $f(n) \geq 0$,
2) non-decreasing property, $\frac{\mathrm{d}f(n)}{\mathrm{d}n} \geq 0$,
3) saturation property, $\frac{\mathrm{d}^2 f(n)}{\mathrm{d}n^2} < 0$.

The first property simply implies that the reward should be non-negative and should be zero, if no one shares genuine information. The second property means that the more UEs share genuine information, the more reward can be gleaned. Finally, the third property suggests that when the amount of genuine information is sufficient, any further increase of the reward remains marginal. In order to construct a general model, in line with the three properties, we opt for the following form of $f(n)$ that has been widely used in economics [27]

$$f(n) = \Gamma - e^{-\mu n}, \text{ if } n \geq 1. \tag{4}$$

In the reward function, μ controls the speed of saturation for the genuine information, while Γ represents the maximum utility of a user, which is a normalized value in the model. This tangible general exponential model can lead to convenient closed-form analysis, which can help us better understand both the model and the mechanism proposed in this paper. In Section V, we will use real-world data to verify this specific form of $f(n)$, which will demonstrate that (4) represents a general formulation of the tangible reward of sharing genuine information among a group of users. In such a case, when there are n UEs sharing genuine information, the utility function of a truthful UE can be simply formulated as:

$$U_{\mathbb{G}}(n) = \Gamma - e^{-\mu n} - c - g(N - n), \tag{5}$$

where again, c is the cost of sharing genuine information (such as privacy jeopardy or battery/resouce depletion cost) and $g(N - n)$ represents the deleterious effect of $N - n$ UEs sharing false information. Note that $g(x)$ has to be a monotonically increasing function of x. Since it represents the deleterious effect of the users sharing false information, it reduces the reward, but indeed, $g(n)$ can be either additive to $f(n)$ in negative form, or a multiplier of $f(n)$ within the range of $[0, 1]$. In this paper, we only consider the additive scenario. Nevertheless, our analytical method and our detailed derivations can be readily extended to the multiplicative scenario. However, we will not discuss the specific formulation of $g(x)$, since it will be cancelled during the derivation of (10). Then, the corresponding utility function can be characterized by

$$U_{\mathbb{F}}(n) = p_a \left(\Gamma - e^{-\mu(n+1)} \right) + (1 - p_a) \left(\Gamma - e^{-\mu n} \right) - g(N - n), \tag{6}$$

where we recall that p_a is the information acquisition probability, while $(n+1)$ in the exponent represents the n pieces of genuine information plus the UE's own genuine information, which has not been shared with other UEs. Below, we will analyze the credibility equilibrium based on this utility model.

B. Evolutionary Credibility Equilibrium

Since the UEs are naturally selfish, they always want to acquire additional genuine information, but potentially without contributing to the information acquisition and sharing owing to its cost. Let us consider the scenario, where given that there is a sufficiently high number of UEs sharing genuine information, some inactive UEs may opt for sharing false information for the sake of reducing their own cost. Gradually, when less and less UEs share genuine information, the utility of each UE may erode to an unsatisfactory level, which in turn, may encourage more and more inactive UEs to abandon their selfish strategy for the sake of increasing their own utility. Eventually, through learning and evaluating the utility improvements in several rounds of interactions with others, each UE can reach a steady-state information sharing strategy, which is a mixed strategy based on ensuring that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

its own utility is maximized. Since a homogeneous model is considered here, a single UE's stable strategy can lead to a stable overall network state, where there is a fixed percentage of UEs sharing genuine information in each time period. We define this stable network state as the information sharing *credibility equilibrium*, which is capable of reflecting the statistical credibility of information across the entire network. This credibility equilibrium is also capable of characterizing a group of selfish UEs' information sharing steady-state without any sophisticated regulation, whilst simultaneously reflecting their tradeoff between the cost of sharing genuine information and the attainable reward.

Apparently, the credibility equilibrium is a result of multiple rounds of multiple UEs' interactions, which cannot be characterized by the traditional Nash equilibrium (NE) of single-shot game interaction. Moreover, the UEs may adopt out-of-NE strategies due to the uncertainty associated with the others' information sharing strategies. In such a case, a robust equilibrium is desired for modelling the credibility equilibrium. The evolutionary equilibrium constitutes a perfect candidate for illustrating such an equilibrium of the periodical dynamic interactions. In evolutionary theory, when some mutants appear, the residential cells can choose either to become a mutant or to maintain their previous status, according to the so-called fitness evaluation including both its own property as well as its interaction with the neighboring cells. Therefore, the dynamics of the periodical information sharing within a group of UEs is quite similar to the dynamics of periodical mutant selection in a set of cells. In the following, we will rely on the evolutionary equilibrium theory [28] for modelling the credibility equilibrium of the information sharing in a cooperative network. Some examples of evolutionary game theory in network modelling include the information diffusion [29], network selection [30], cognitive radio networks [31], [32] adaptive filtering networks [33], P2P steaming [34] and social networks [35].

Let us denote the credibility equilibrium as $P_{\mathbb{G}}^*$, which represents the steady-state probability of a UE sharing its genuine information across the network. According to the definition of the evolutionary game and the evolutionarily stable strategy [28], $P_{\mathbb{G}}^*$ should satisfy both the NE property and stability property described in the following definition.

Definition 1: In an information sharing network, the credibility equilibrium $P_{\mathbb{G}}^*$ should satisfy

1) NE property: $U_i(P_{\mathbb{G}}, P_{\mathbb{G}}^*) \leq U_i(P_{\mathbb{G}}^*, P_{\mathbb{G}}^*)$,
2) stability property: if $U_i(P_{\mathbb{G}}, P_{\mathbb{G}}^*) = U_i(P_{\mathbb{G}}^*, P_{\mathbb{G}}^*)$, $U_i(P_{\mathbb{G}}, P_{\mathbb{G}}) < U_i(P_{\mathbb{G}}^*, P_{\mathbb{G}})$,

where $U_i(x, y)$ is UE i's utility when it adopts strategy x while other UEs adopts strategy y.

We can see that the first condition is the NE condition, suggesting that no other strategy is capable of unilaterally swaying the UEs into deviating from $P_{\mathbb{G}}^*$; while the second condition guarantees the stability of the strategy, implying that even if there exists some strategy $P_{\mathbb{G}}$ that performs equally well as $P_{\mathbb{G}}^*$, $P_{\mathbb{G}}^*$ must be a better response-action to $P_{\mathbb{G}}$ than $P_{\mathbb{G}}$ itself. Recall that our target in this section is to calculate the steady-state equilibrium $P_{\mathbb{G}}^*$. To achieve that, we first have to find an appropriate representation of the UEs' interaction dynamics and then analyze the converged solutions of the dynamics, which can be the specific candidates of the credibility equilibrium.

Since all UEs are uncertain about the other UEs' information sharing strategies and utilities, each UE has to try different tentative strategies for improving its own utility throughout the different rounds of interactions and learn by inference from the interactions using the intuitive methodology of understanding-by-building. According to evolutionary game theory, the following differential equation, called replicator dynamics, can be used for modelling the corresponding interactive dynamics within a group of UEs [28],

$$\dot{P}_{\mathbb{G}} = \alpha P_{\mathbb{G}} \left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}(P_{\mathbb{G}}) \right], \tag{7}$$

where $\dot{P}_{\mathbb{G}}$ is a mutated variant of $P_{\mathbb{G}}$, α is a positive coefficient, $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})$ is the average utility of these specific UEs, which share genuine information and $\bar{U}(P_{\mathbb{G}})$ is the average utility of all UEs in the entire network. We can see that if the sharing of genuine information can lead to a higher utility than the average level, the percentage of altruistic UEs sharing genuine information will increase and the rate of increase formulated as $\dot{P}_{\mathbb{G}}/P_{\mathbb{G}}$ is proportional to the difference between $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})$ and $\bar{U}(P_{\mathbb{G}})$. When each UE shares genuine information with a probability of $P_{\mathbb{G}}$, the average utility $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})$ of sharing genuine information can be calculated according to (5) as follows

$$\begin{aligned} \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) = \sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^n (1-P_{\mathbb{G}})^{N-1-n} \\ \left(\Gamma - e^{-\mu(n+1)} - c - g(N-n) \right), \end{aligned} \tag{8}$$

where $\binom{N-1}{n} P_{\mathbb{G}}^n (1-P_{\mathbb{G}})^{N-1-n}$ is the configuration probability that there are n UEs among the $(N-1)$ other UEs sharing genuine information. Similarly, the average utility $\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})$ of sharing false information can be calculated according to (6) as follows

$$\begin{aligned} \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) = \sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^n (1-P_{\mathbb{G}})^{N-1-n} \\ \cdot \left[p_a \left(\Gamma - e^{-\mu(n+1)} \right) + (1-p_a) \left(\Gamma - e^{-\mu n} \right) - g(N-n) \right]. \end{aligned} \tag{9}$$

Thus, the resultant average utility of the entire network can be calculated by

$$\begin{aligned} \bar{U}(P_{\mathbb{G}}) &= P_{\mathbb{G}} \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) + (1-P_{\mathbb{G}}) \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) \\ &= \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) + P_{\mathbb{G}} \left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) \right]. \end{aligned} \tag{10}$$

Furthermore, by combining (7) and (10), as well as by using

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

$$P^*_{\mathbb{G}} = \begin{cases} 0, & \frac{c}{1-p_a} \geq 1-e^{-\mu}, \\ \left(1-\frac{c}{(1-p_a)(1-e^{-\mu})}\right)^{\frac{1}{N-1}} / \left(1-e^{-\mu}\right), & 1-e^{-\mu} < \frac{c}{1-p_a} < e^{-\mu(N-1)} - e^{-\mu N}, \\ 1, & 0 \leq \frac{c}{1-p_a} \leq e^{-\mu(N-1)} - e^{-\mu N}. \end{cases} \tag{12}$$

(8) and (9), we arrive at:

$$\begin{aligned} \dot{P}_{\mathbb{G}} &= \alpha P_{\mathbb{G}}(1-P_{\mathbb{G}}) \left(\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right) \\ &= \alpha P_{\mathbb{G}}(1-P_{\mathbb{G}}) \Bigg(\sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^n (1-P_{\mathbb{G}})^{N-1-n} \\ &\quad \left[(1-p_a)\left(e^{-\mu n} - e^{-\mu(n+1)}\right)\right] - c \Bigg) \\ &= \alpha P_{\mathbb{G}}(1-P_{\mathbb{G}}) \Bigg((1-p_a) \Bigg(1 - e^{-\mu} \Bigg) \\ &\quad \sum_{n=0}^{N-1} \binom{N-1}{n} \left(e^{-\mu} P_{\mathbb{G}}\right)^n (1-P_{\mathbb{G}})^{N-1-n} - c \Bigg) \\ &= \alpha P_{\mathbb{G}}(1-P_{\mathbb{G}}) \Big((1-p_a) \left(1-e^{-\mu}\right) \\ &\quad \left(1 - \left(1-e^{-\mu}\right) P_{\mathbb{G}}\right)^{N-1} - c \Big). \end{aligned} \tag{11}$$

When each UE reaches the steady state, any changes in the information sharing strategy should converge to 0, i.e. we have $\dot{P}_{\mathbb{G}} = 0$. By equating the right-hand-side of (11) to 0, we can have three solutions, $P^{*1}_{\mathbb{G}} = 0$, $P^{*2}_{\mathbb{G}} = 1$ or $P^{*3}_{\mathbb{G}} = \left(1-\frac{c}{(1-p_a)(1-e^{-\mu})}\right)^{\frac{1}{N-1}} / (1-e^{-\mu})$. Apparently, the first two solutions, namely $P^{*1}_{\mathbb{G}} = 0$ and $P^{*2}_{\mathbb{G}} = 1$ represent the boundaries, which correspond to the scenarios of having no one sharing genuine information and having all UEs sharing genuine information, respectively. While the third solution $P^{*3}_{\mathbb{G}}$ should be within the interval $(0, 1)$, i.e. a portion of UEs shares genuine information. Note that these three solutions are not necessarily evolutionary credibility equilibrium, which should satisfy the aforementioned NE property and stability property. The following theorem shows the conditions of being evolutionary credibility equilibrium for each solution.

Theorem 1: The N-UE evolutionary credibility equilibrium when the information acquisition probability p_a is constant can be written as (12).

Proof: Please see the appendix.

Remarks: In (12), the term $\frac{c}{1-p_a}$ represents the expected cost of a specific UE sharing genuine information. The physical interpretation of the condition for $P^*_{\mathbb{G}} = 0$ is that the cost of sharing genuine information for each UE is higher than the reward each UE can obtain, when there is a single UE sharing genuine information. This implies that if the network is currently in the undesirable state, when no UE shares genuine information, no single UE has the incentive to unilaterally change its strategy, since again the reward is lower than the cost. By contrast, the physical meaning of the condition $P^*_{\mathbb{G}} = 1$ is that, the information sharing cost is lower than the increase of the reward, when the network's status evolves from $(N-1)$ UEs sharing genuine information to all UEs. This means that if the network is currently in the state of all UE sharing genuine information, no single UE has the incentive to unilaterally change its strategy, since its reward is higher than its cost. When the cost is in the middle of the range corresponding to the second scenario of (12), the probability of each UE sharing genuine information converges to a fixed value, which means that there is a fixed percentage of UEs sharing genuine information in each round of interaction. This percentage is inversely proportional to the information sharing cost c. Therefore, given the network settings specified in terms of the parameters μ and Γ, in order to encourage more UEs to share genuine information and hence to enhance the final steady-state credibility equilibrium, the service provider should control the information sharing cost in order to satisfy the aforementioned third condition of (12).

V. Heterogeneous Model: How Credit can Help

In the previous section, we have studied the spontaneous credibility equilibrium of a homogeneous model, where no sophisticated regulation was carried out. By contrast, this section will consider a heterogeneous information sharing model, where all the UEs have different mixed strategies and each UE adopts different strategies for interacting with different UEs. More importantly, a credit mechanism is invoked in this model for the sake of encouraging more UEs to share genuine information across the network. However, for the sake of smoothly increasing the grade of sophistication, we will first study the spontaneous information credibility equilibrium of this heterogeneous model without any credit mechanism. Since in contrast to the homogeneous model, the heterogeneous setting does not lend itself to a closed-form credibility equilibrium formulation, we will rely on the powerful reinforcement learning algorithm for numerically finding its solution. Then, we propose a credit mechanism, which can ensure that the UEs actively share genuine information with each other.

A. Utility Formulation

Similar to the homogeneous model, UEs in the heterogeneous setting also share their information periodically with each other for making decisions, following one of the two possible strategies $\mathbb{G}$ and $\mathbb{F}$, i.e. either sharing genuine information or sharing false information. Again, in contrast to the homogeneous model, each UE adopts different information sharing strategies with respect to different UEs. Let us denote UE i's strategy vector $\mathbf{s}_i$ and action vector $\mathbf{a}_i$ as

$$\mathbf{s}_i = [s_{i1}, s_{i1}, \ldots, s_{iN}], \tag{13}$$

$$\mathbf{a}_i = [a_{i1}, a_{i1}, \ldots, a_{iN}], \tag{14}$$

where $0 \leq s_{ij} \leq 1$ represents the probability of UE i sharing genuine information with UE j, $a_{ij} = \mathbb{G}$ or $\mathbb{F}$ represents UE i's specific action instantiated by s_{ij} corresponding to either sharing genuine or false information with UE j, $s_{ii} \equiv 1$ and $a_{ii} \equiv \mathbb{G}$ represents that UE i is always aware of its own

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

genuine information. Moreover, in order to formulate different strategies for the different UEs, each UE has to evaluate the shared information in a pair-wise manner, instead of relying on a population-based homogeneous model. In the homogeneous model, since all users are considered to be identical, the utility function is defined among a group of identical users, which can help reveal the global equilibrium of the whole group. By contrast, when it comes to the heterogenous model, since all users can adopt different strategies in conjunction with different utilities and the interaction also becomes a one-to-one action, theoretically the global equilibrium cannot be reached and thus the utility function is defined between two users. This requires each UE to immediately process the information after interacting with a specific UE and to evaluate the resultant information utility, which can be used for updating its strategy with respect to this specific UE. Let us define UE i's utility vector as

$$\mathbf{U}_i = [U_{i1}, U_{i1}, \ldots, U_{iN}], \tag{15}$$

where U_{ij} represents UE i's utility after exchanging information with UE j. According to the reward function defined in (4), the utility $U_{ij/ji}(\mathbb{G}, \mathbb{G})$ can be formulated as:

$$U_{ij/ji}(\mathbb{G}, \mathbb{G}) = \Gamma - e^{-2\mu} - c_{ij/ji}, \tag{16}$$

where $U_{ij}(x, y)$ represents UE i's utility, when it adopts strategy x, while UE j adopts strategy y, the information sharing costs c_{ij} and c_{ji} are different for the different UEs in this heterogenous model. Furthermore, the utilities $U_{ij/ji}(\mathbb{G}, \mathbb{F})$ and $U_{ij/ji}(\mathbb{F}, \mathbb{G})$ are expressed as:

$$U_{ij/ji}(\mathbb{G}, \mathbb{F}) = \Gamma - e^{-\mu} - c_{ij/ji} - g, \tag{17}$$

$$U_{ij/ji}(\mathbb{F}, \mathbb{G}) = \Gamma - e^{-2\mu}, \tag{18}$$

where g is the delirious effect of the false information. It is assumed that the UE always acquires new information in each time period, since it is possible that no others share genuine information with it in this heterogeneous model. For the utility $U_{ij/ji}(\mathbb{F}, \mathbb{F})$, we have

$$U_{ij/ji}(\mathbb{F}, \mathbb{F}) = \Gamma - e^{-\mu} - g. \tag{19}$$

B. Spontaneous Credibility Equilibrium: Reinforcement Learning

When considering the spontaneous information sharing process from an individual UE's perspective, the network of all other UEs can be regarded as an external environment. The UE makes decisions and carries out its information sharing actions for maximizing its utility in this environment, which is a dynamic environment since all other UEs' strategies are also adjusted as a result of their interactions. Generally, each UE learns from its interactions with the environment and adapts to the environment by adjusting its strategies for the sake of increasing its utility attained. Reinforcement learning is a powerful tool conceived for tackling adaptive environment-learning and decision-making problems [36]. In contrast to the supervised learning associated with correct input/output pairs, reinforcement learning was inspired by behaviorial psychology and aims for maintaining a certain on-line performance by striking a balance between the exploration of an uncharted territory and the exploitation of current knowledge. It is capable of learning an unknown environment's statistics as well as of taking actions in the environment so as to maximize the cumulative reward, where the environment itself may be changed by the agent's actions. Reinforcement learning has been widely adopted in communications and networks [37], in control [38], in finance and economics [39], as well as in social science [40]. Specifically, Xiao *et. al.* has initiatively applied the reinforcement learning in network security modeling [41]-[43].

In this reinforcement learning process, each UE should periodically update its information sharing strategy $\mathbf{s}_i$ through learning from its interactions with others, while the quantitative characterization of its interactions is provided by the utility gleaned from its interaction with others. Hence the challenge is how to update UE i's strategy $\mathbf{s}_i$ according to the utility $\mathbf{U}_i$ inferred from the current interaction, as well as that obtained during its past experiences. The reinforcement learning model introduced the concept of *perception* based on the accumulated utilities associated with its actions, which records both the results of all the past interactions as well as the new interaction's result, corresponding to the exploitation of past knowledge and to the exploration of the new environment. Let us define UE i's perception matrix $\mathbf{P}_i$ as

$$\mathbf{P}_i = [\mathbf{p}_{i1}, \mathbf{p}_{i2}, \ldots, \mathbf{p}_{iN}], \tag{20}$$

$$\mathbf{p}_{ij} = [p_{ij\mathbb{G}}, p_{ij\mathbb{F}}]^T, \tag{21}$$

where $p_{ij\mathbb{G}}$ is UE i's perception after sharing its genuine information with UE j, and $p_{ij\mathbb{F}}$ is that of sharing false information. At the end of each interaction, UE i first evaluates its own utility and then uses this utility value for adjusting its perception associated with the specific action just adopted and keeps the perception of the complimentary action unchanged, which can be expressed by

$$p_{ij\mathbb{G}}^{t+1} = \begin{cases} (1-\epsilon_i^t)p_{ij\mathbb{G}}^t + \epsilon_i^t U_{ij}^t, & \text{if } a_{ij}^t = \mathbb{G}, \\ p_{ij\mathbb{G}}^t, & \text{otherwise}, \end{cases} \tag{22}$$

$$p_{ij\mathbb{F}}^{t+1} = \begin{cases} (1-\epsilon_i^t)p_{ij\mathbb{F}}^t + \epsilon_i^t U_{ij}^t, & \text{if } a_{ij}^t = \mathbb{F}, \\ p_{ij\mathbb{F}}^t, & \text{otherwise}, \end{cases} \tag{23}$$

where the superscript t represents the time period, U_{ij}^t is UE i's utility gained by exchanging information with UE j at time instant t, while ϵ_i^t represents a sequence of averaging factors controlling the discounting rate associated with $\sum_t \epsilon_i^t = \infty$ and $\sum_t (\epsilon_i^t)^2 < \infty$. Explicitly, physical interpretation of this discounting rate is the weight of the newly acquired utility in evaluating the perception. We assume that UE i has no additional information about the other UEs' actions or utilities during the interaction, which is indeed typical case in practical scenarios.

After updating the perception $\mathbf{p}_{ij}$, UE i can exploit it for constructing its next information sharing strategy with respect to UE j, which will be used later for generating its information sharing action as regards to UE j. A beneficial method is for UE i to maintain a specific proven action that maximizes its expected utility. Explicitly, if we have $p_{ij\mathbb{G}} > p_{ij\mathbb{F}}$, then UE i would share genuine information with UE j, but it can only infer the resultant perception $p_{ij\mathbb{G}}$ in the next round and so

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

on. However, this conservative method may incur the risk of inadequately exploring all other legitimate actions. Therefore, the construction of both UE i's reputation vector as well as of its mixed strategy and of its action is associated with an exploration method. A commonly adopted simple exploration method is to pursue the specific action associated with the best expected perception by default, but also occasionally choose the other actions at random with a probability of p. This method usually starts with a large value of p to encourage suffeciently diverse initial exploration, which is then gradually reduced in the vicinity of the optimum to expedite convergence. In contrast to this simple method, the more sophisticated *Boltzmann* exploration exhibiting a better performance is also widely used in reinforcement learning [36], where the already available perception is used for probabilistically choosing an action according to the Boltzmann distribution. Explicitly, in our model UE i's strategy with respect to UE j can be updated using the already available perception by invoking:

$$s_{ij}^t = \frac{e^{\xi_j^t p_{ij\mathbb{G}}^t}}{e^{\xi_j^t p_{ij\mathbb{G}}^t} + e^{\xi_j^t p_{ij\mathbb{F}}^t}}, \tag{24}$$

where the positive coefficient ξ_j^t controls the exploration level, with $\xi_j^t \to 0$ leading to a 50% probability for both $\mathbb{G}$ and $\mathbb{F}$, while for $\xi_j^t \to \infty$ the action would concentrate only on the pure $\mathbb{G}$ or $\mathbb{F}$ strategy, depending on which leads to a higher perception.

To summarize, the reinforcement learning based credibility equilibrium learning process can be interpreted as a process in which each UE simultaneously probes the pure strategies $\mathbb{G}$ and $\mathbb{F}$ with respect to the other UEs in order to learn about its own bilateral utilities as well as perceptions, and updates its estimation regarding the other UEs' reputation as well as adjusting its own interaction in the next round accordingly by using the accumulated perception. The iteration evolving from $\mathbf{p}_{ij}^t$ to $\mathbf{p}_{ij}^{t+1}$ can be illustrated by a chain of elementary steps: the initial perception gives rise to a random mixed strategy that determines the specific action by taking this specific action, the utility is evaluated and then the perception can be updated in the next round, and so on. The iterations can be simply represented by $\mathbf{p}_{ij}^t \to s_{ij}^t \to a_{ij}^t \to U_{ij}^t \to \mathbf{p}_{ij}^{t+1}$, as depicted in Fig. 3. Following this reinforcement learning procedure, the UEs are expected to converge to a stable credibility equilibrium in the network.

The dynamics formulated in (22) can be written in a recursive vectorial form as follows

$$\mathbf{p}_{ij}^{t+1} - \mathbf{p}_{ij}^t = \epsilon^t(\mathbf{q}_{ij}^t - \mathbf{p}_{ij}^t), \tag{25}$$

where $\mathbf{q}_{ij}^t = [q_{ij\mathbb{G}}^t, q_{ij\mathbb{F}}^t]$ can be calculated by

$$q_{ij\mathbb{G}}^t = \begin{cases} U_{ij}^t, & \text{if } a_{ij}^t = \mathbb{G}, \\ p_{ij\mathbb{G}}^t, & \text{otherwise}, \end{cases} \tag{26}$$

$$q_{ij\mathbb{F}}^t = \begin{cases} U_{ij}^t, & \text{if } a_{ij}^t = \mathbb{F}, \\ p_{ij\mathbb{F}}^t, & \text{otherwise}. \end{cases} \tag{27}$$

If both sides of (25) are divided by a sufficiently small coefficient ϵ^t, the recursive equation can be approximated by a differential equation of continuous-time deterministic averaged dynamics, i.e.,

$$\dot{\mathbf{P}}_i = \mathbb{E}(\mathbf{Q}_i|\mathbf{P}_i) - \mathbf{P}_i, \tag{28}$$

where $\mathbb{E}(\cdot|\mathbf{P}_i)$ is the expectation conditioned on the action a_{ij}, which is determined by $\mathbf{p}_{ij}$ as in (24). We can see that when the reinforcement learning process becomes converged, i.e. we have $\dot{\mathbf{P}}_i = 0$, then the perception should become equivalent to the expected utility a UE can obtain. The convergence of (28) is closely related to the exploration level ξ^t in (24), since an aggressive exploration may lead to rapid convergence. The following theorem gives a sufficient condition that ensures the convergence of the proposed algorithm.

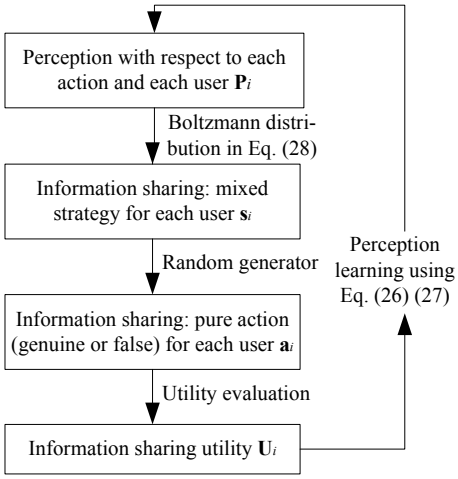

Fig. 3. Reinforcement learning chain of the homogeneous model.

Theorem 2: The condition of convergence for the proposed reinforcement learning scheme is that the exploration level ξ should satisfy

$$\xi \leq \frac{1}{2(e^{-\mu} - e^{-2\mu} + g)}. \tag{29}$$

Proof: Please see the appendix.

Theorem 2 shows the convergence condition of the reinforcement learning algorithm. A larger ξ accelerates the convergence but it should remain lower than some threshold, which is related to the utility function. In practical scenarios, ξ is chosen by the individual user, since the algorithm is fully distributed. In contrast to the homogeneous model, which lent itself to closed-form analysis of the steady-state equilibrium, the dynamics of the heterogeneous information sharing model in (28) cannot be characterized by a closed-form expression. Hence it will be characterized by simulations in Section V.

C. Credit Mechanism

In human networks, everyone has a certain reputation or credit record, which is generated and updated according to one's behavior in social networks. When a person interacts with another one having a good credit/reputation, they would be inclined to maintain contact. By contrast, nobody likes to cooperate with people carrying a bad reputation. Similarly, in a cooperative network, each UE is capable of constructing a credit-belief through its interactions with other UEs. Moreover, each UE determines, whether to share its information with

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

another UE according to that UE's credit. When a high-credit UE discovers another UE with a credit lower than some threshold, a high-credit UE would be uninclined to share any information with the low-credit UE. In such a case, through rounds of interactions, each UE's credit can be gradually inferred by the observation of shared information, leading to the fact that the UEs having a low credit would obtain less and less genuine information and eventually they would have to change their information sharing strategy in order to improve their reputation. In this section, we will propose a credit mechanism and investigate how assigning credits can help enhance the steady-state credibility equilibrium.

Similar to human networks, each of the UEs in a cooperative network has a credit value that is commensurate with its past behavior and also determines its future behavior. Let us define the UEs' reputation vector as

$$\mathbf{r} = [r_1, r_2, \ldots, r_N], \tag{30}$$

where $0 \leq r_i \leq 1$ represents UE i's credit/reputation in the network. In human networks, a person's future behavior tends to be consistent with his/her past reputation, regardless of the credit-level of the other persons he/she is interacting with. In other words, a "tit-for-tat" behavior cannot help a person to maintain his/her reputation. However, a reputable person may opt for avoiding contact with the ones having a bad reputation. Similarly, a UE's information sharing strategy in the cooperative network should also be consistent with its past reputation and should be independent of the reputation of the UEs it is confronted with. Therefore, UE i' mixed information sharing strategy should be independent of that of all others and ought to be proportional to its own credit, i.e. we have:

$$\mathbf{s}_i = [r_i, r_i, \ldots, r_i]. \tag{31}$$

Nevertheless, when UE i has the knowledge of UE j's credit value through rounds of interactions, it can determine, whether to share information with UE j in the future. Let us define UE i's interaction probability/decision with respect to the others as

$$\boldsymbol{\rho}_i = [\rho_{i1}, \rho_{i2}, \ldots, \rho_{iN}], \quad \mathbf{d}_i = [d_{i1}, d_{i2}, \ldots, d_{iN}], \tag{32}$$

where $0 \leq \rho_{ij} \leq 1$ represents UE i's probability of sharing information with UE j, regardless, whether this is genuine or false information, and $d_{ij} = 0$ or 1 represents whether to share information with UE j in a specific round of interactions. In such a case, at the beginning of each time period, UE i determines both

- the information sharing decision $\mathbf{d}_i$, namely whether to share information with a specific UE, according to its interaction probability $\boldsymbol{\rho}_i$;
- and the information sharing action a_i with respect to the UEs its has chosen to share information with, i.e. whether to share genuine or false information, according to its strategy $\mathbf{s}_i$ and reputation r_i.

Meanwhile, after several rounds of interactions, UE i should update its interaction probability $\boldsymbol{\rho}_i$ according to its past experience with others. It is expected that following these alternating decision making and learning phases, the UEs having a bad reputation would received less and less information from the others and hence they would have to increase their credit value by actively sharing genuine information hereafter.

We still have to resolve the problem of how each UE generates its interaction probability vector by learning the behaviors of others. Here, we can also use reinforcement learning for solving this problem, as in the previous subsection. Let us define UE i's perception of the others' behavior as ψ_i, where

$$\boldsymbol{\Psi}_i = [\psi_{i1}, \psi_{i2}, \ldots, \psi_{iN}]. \tag{33}$$

The UE i's perception ψ_{i1} regarding UE j's reputation can be learned through evaluating the information utility gleaned from UE j, which can be constructed by

$$\psi_{ij}^{t+1} = \begin{cases} (1-\epsilon_i^t)\psi_{ij}^t + \epsilon_i^t U_{ij}^t, & \text{if } d_{ij}^t = 1, \\ \psi_{ij}^t, & \text{if } d_{ij}^t = 0. \end{cases} \tag{34}$$

The resultant perception can then be used for generating UE i's interaction probability. Apparently, the higher utility UE i can infer through exchanging information with UE j, the higher the interaction probability ρ_{ij} should be, indicating a proportional relationship between ρ_{ij} and ψ_{ij}. Here, we adopt a normalized performance evaluation method as follows

$$\rho_{ij}^t = \frac{e^{\xi_i^t \psi_{ij}^t}}{\max\limits_j \{e^{\xi_i^t \psi_{ij}^t}, \forall j\}}, \tag{35}$$

where the physical meaning of (35) is that UE i always continues its interaction with the specific UE yielding the highest utility. Furthermore, it uses this highest utility as a reference, when UE i determines its probability of interaction with others. In general, we can summarize the credit-based reinforcement learning procedure as a chain

$$\begin{array}{ccccccc} \boldsymbol{\Psi}_i^t & \rightarrow & \boldsymbol{\rho}_i^t & \rightarrow & \mathbf{d}_i^t & \rightarrow & \mathbf{U}_i^t \rightarrow \boldsymbol{\Psi}_i^{t+1}, \\ & & \downarrow & & \uparrow & & \\ & & r_i^t & \rightarrow & \mathbf{s}_i^t & \rightarrow & a_i^t \end{array} \tag{36}$$

where the arrow between $\boldsymbol{\rho}_i^t$ and r_i^t implies that when a UE discovers that the number of others sharing information with it falls below some threshold, the UE would consider to increase its own credit value in order to enhance its reputation by sharing more genuine information with the others. This credit mechanism is summarized in Algorithm 1.

VI. Simulation Results

In this section, we first use a pair of data traces to justify the choice of the reward function $f(n)$ defined in (4). The essence of the utility function is the reward function $f(n)$, whilst the cost is constant. Therefore, we only characterize the reward function using real-world data traces. In the first data trace, we evaluate the relationship between the number of users providing genuine rating information and the user's utility, which is defined as the recommendation quality. In the second data trace, we evaluate the relationship between the number of users providing genuine private data information and the user's utility. Finally, we conduct simulations to verify both the proposed algorithms and our theoretical analysis.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

Algorithm 1 Credit mechanism for information sharing.

for each UE i **do**
  /********** *Initialization* **********/
  Initialize UE i's credit value r_i^0 and credit adjustment step size Δr_i.
  Initialize UE i's perception $\mathbf{\Psi}_i^0 = \mathbf{0}$.
  Initialize the number of interactive UEs i $n_i^0 = 0$ and the threshold n_T.
  Setup the learning speed ϵ_i, the exploration level ξ_i and the tolerance ζ.
  /********** *Information sharing interaction* **********/
  while $\sum_j \left(\psi_{ij}^t - \psi_{ij}^{(t-1)}\right)^2 \geq \zeta$ **do**
    Calculate $\boldsymbol{\rho}_i^t$ by $\rho_{ij}^t = e^{\xi_i^t \psi_{ij}^t} / \max\{e^{\xi_i^t \psi_{ij}^t}, \forall j\}$.
    Determine $\mathbf{d}_i^t$ using random number generator $\mathbf{rand}(\boldsymbol{\rho}_i^t)$.
    Set $s_i^t = r_i^t$ and determine a_i^t using $\mathbf{rand}(s_i^t)$.
    Information sharing, evaluate the utility $\mathbf{U}_i^t$ and store n_i^t.
    Update UE i's perception $\mathbf{\Psi}_i^t$ using (34).
    $t = t + 1$.
  end while
  /********** *Reputation adjustment* **********/
  if $\frac{1}{t}\sum_t n_i^t < n_T$ **then**
    $r_i = r_i + \Delta r_i$.
  end if
end for

A. Reward Function Verification

1) Recommender Systems: In recommender systems, the more users share genuine ratings, the more reliable the recommendation becomes, which can be defined as the utility for the users. More specifically, we utilize the Jester data set [44] to evaluate this relationship, which contains about 4.1 million ratings of 100 jokes from 73,421 users. Since the "genuine truth" is required for evaluating the recommendation quality, we consider 720,000 ratings from the 7,200 users who have rated all the 100 jokes. The user-item matrix $\mathbf{R} = [r_{ij}]_{7200\times 100}$ contains non-zero elements, and each row of $\mathbf{R}$ is treated as the corresponding user's interest vector. Based on the matrix $\mathbf{R}$, we construct a group of rating matrices $\{\mathbf{R}_{i,k}\}$ with false rating information sharing, by manipulating $\rho_k\%$ of the ratings matrix $\{\mathbf{R}_{i,k}\}$ to "0", where $\rho_k \in \{0, 1, \cdots, 90\}$. Then, we apply the similarity-based collaborative filtering algorithm to recover the original rating matrix as follows: a recommendation vector $\hat{\mathbf{r}}_i = (\hat{r}_{i1}, \cdots, \hat{r}_{i100})$ is computed for user i, where $\hat{r}_{ij}$ is defined as

$$\hat{r}_{ij} = \begin{cases} r_{ij}, & \text{if } r_{ij} \neq 0, \\ h_{ij}^{CF}(\mathbf{R}), & \text{if } r_{ij} = 0, \end{cases} \tag{37}$$

where the recovered rating $h_{ij}^{CF}(\mathbf{R})$ is defined as:

$$h_{ij}^{CF}(\mathbf{R}) = \frac{\sum\limits_{k \in Neighbour(i)} r_{kj} F_{sim}(i,k)}{\sum\limits_{k \in Neighbour(i)} F_{sim}(i,k)}, \tag{38}$$

with $F_{sim}(i,k)$ representing the similarity between user i and user k, while $Neighbour(i)$ representing the set of users who are most similar to user i. The similarity $F_{sim}(i,k)$ is measured in terms of the vector cosine similarity metric. As mentioned above, the users' utility is defined as the recommendation quality, which is calculated as follows. Let $\mathbf{p}_i = (p_{i1}, \cdots, p_{iM})$ denote user i's interest, where p_{ij} represents user i's true preference for item s_j $(j = 1, \cdots, M)$. The quality of $\hat{\mathbf{r}}_i$ is evaluated by a user-specific function $g_i : \mathbb{R}^M \to \mathbb{R}$, which is defined as

$$g_i(\hat{\mathbf{r}}_i) = 1 - \frac{\sqrt{\sum\limits_{j=1}^{M} (\hat{r}_{ij} - p_{ij})^2}}{r_{\max}\sqrt{M}}. \tag{39}$$

A large $g_i(\hat{\mathbf{r}}_i)$ implies having substantial similarity between $\mathbf{r}_i$ and $\mathbf{p}_i$, namely a high recommendation quality and a high utility. Finally, let us define $\sigma_R = 100 - \rho_k$, which is the percentage of users sharing genuine ratings. In such a case, for each σ_R, we can evaluate the corresponding utility of each user by implementing the collaborative filtering algorithm of (37-38) and the recommendation quality evaluation of (39). Fig. 4 shows the relationship between the number of users sharing genuine rating and the utility, from which we can see that the relationship can be perfectly reflected by the reward function $f(n)$ defined in (4) by using the curve fitting toolbox provided in MATLAB.

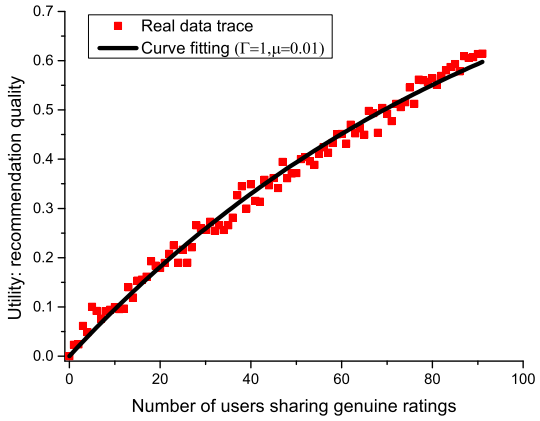

Fig. 4. Recommender system.

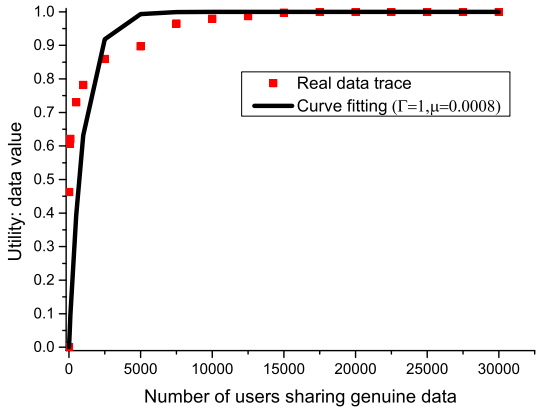

Fig. 5. Data mining.

2) Data Mining: In data mining, the more users share genuine data, the more value can be gleaned from it, which can be defined as the utility of the users. In this context we utilize the Adult data set [45], which consists of 32 561 records from a census database, and each record consists of 15 attributes. After removing the records with missing values,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

we use the remaining 30 162 records for our experiments. In data mining, anonymization is a mandatory procedure to protect the users' information privacy. We developed a Java program based on the open source anonymization framework ARX of [46], which supports different types of privacy criteria and provides multiple methods for quantifying the value of information after anonymization and mining [47]. Here, we opted for realying on k-anonymity as the privacy criterion and the information value loss is evaluated by the recommended default measure $Loss$ [48]. The value of $Loss$ ranges from 0 to 1, where a large value indicates a high information loss. As mentioned above, our objective is to verify that the value of the information increases with the number of users sharing genuine data records. To achieve this objective, we randomly select N genuine records in the Adult dataset and manipulate the remaining $30162 - N$ records to render tham false in order to simulate the information sharing process. Then, we run the anonymization program on these data sets respectively and record the corresponding information value $1 - Loss$, which is defined as the utility of the users. We observe from the experimental results shown in Fig. 5 that again, this relationship is perfectly fitted by the reward function $f(n)$ defined in (4).

B. Homogeneous Model

To verify the credibility equilibrium of the homogeneous model, we simulate a group of UEs' information sharing inclinations in the network. It can be seen that the updating of $P_{\mathbb{G}}$ in (7) requires global information, such as the average utility of sharing genuine information and the average utility of all UEs. Since an individual UE has no access to such global information in the distributed network, a beneficial strategy for each UE is to adopt a specific mixed strategy for a certain period, during which the utilities may be circulated for constructing an average value from each UE's own perspective. Let us first discretize (7) as follows

$$P_{\mathbb{G}}^{t+1} = P_{\mathbb{G}}^{t} + \alpha \left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{t}) - \bar{U}(P_{\mathbb{G}}^{t})\right]. \tag{40}$$

Assuming that UE i adheres to its mixed strategy $P_{\mathbb{G}}^{t}$ for a while to estimate both $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{t})$ and $\bar{U}(P_{\mathbb{G}}^{t})$, let us define an indicator function $\mathbf{1}_i(t,k)$ as

$$\mathbf{1}_i(t,k) = \begin{cases} 1, & \text{if UE } i \text{ shares genuine information,} \\ 0, & \text{if UE } i \text{ shares false information.} \end{cases} \tag{41}$$

Then, after several rounds of interactions, the approximated $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{t})$ and $\bar{U}(P_{\mathbb{G}}^{t})$ can be calculated by

$$\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{t}) = \frac{\sum_k U_i(t,k)\mathbf{1}_i(t,k)}{\sum_k \mathbf{1}_i(t,k)}, \tag{42}$$

$$\bar{U}(P_{\mathbb{G}}^{t}) = \frac{1}{|k|}\sum_k U_i(t,k), \tag{43}$$

where $U_i(t,k)$ is the utility obtained and evaluated by UE i at time instant k.

We simulate such an iterative interaction process among $N = 20$ homogeneous UEs in conjunction with setting the utility function as $\Gamma = 1$ and $\mu = 0.1$, as well as relying on the step size $\alpha = 1$. In Fig. 6, we show the simulated dynamics of the information sharing process, when the sharing cost of genuine information is $c = 0.01$, where the y-axis is the probability of an individual UE sharing genuine information. The simulations have been conducted for different starting points, i.e. for different initial genuine information sharing probability of $P_{\mathbb{G}}^{0}$. We can see that the final converged credibility equilibrium is independent of the initial point, which is purely determined by the dynamics and by the fixed point formulated in (7). In Fig. 7, we show the

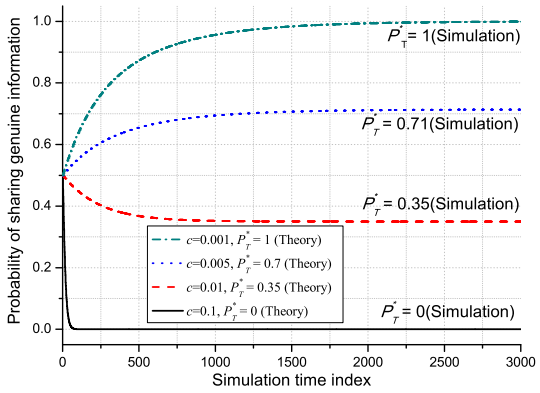

Fig. 6. Impact of starting point on the steady-state credibility equilibrium learning process.

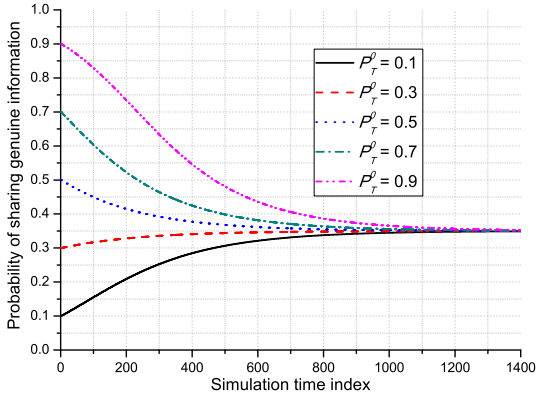

Fig. 7. Impact of information sharing cost on the steady-state credibility equilibrium.

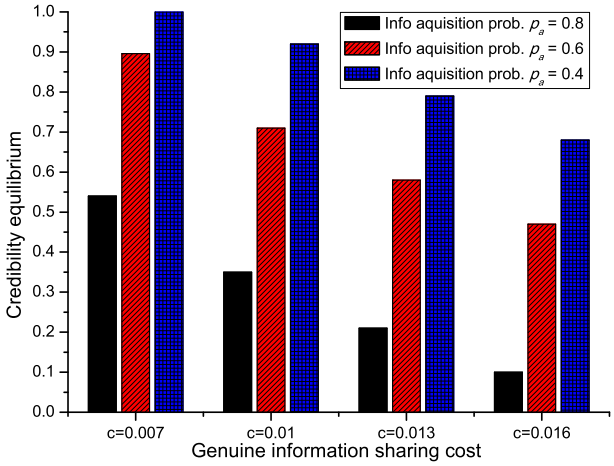

Fig. 8. Impact of information acquisition probability cost on the specific value of the steady-state credibility equilibrium.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

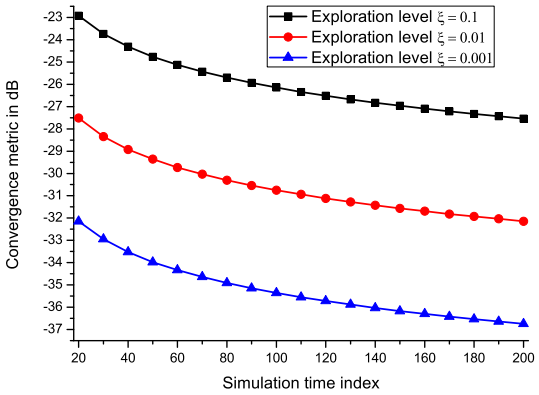

Fig. 9. Convergence performance of the reinforcement learning.

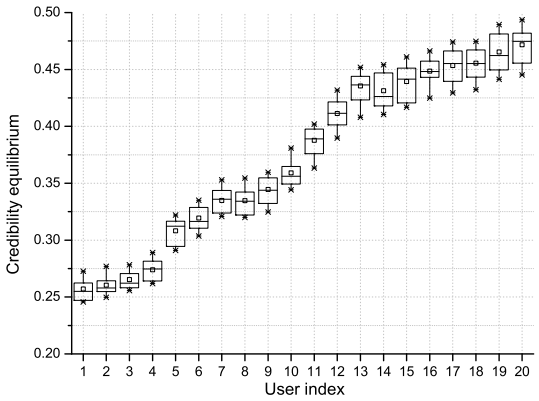

Fig. 10. Credibility equilibrium of the heterogeneous model.

impact of the genuine information sharing cost on the final value of the steady-state credibility equilibrium, where the four lines correspond to the four costs of $c = 0.001,\ 0.005,\ 0.01$ and 0.1, respectively. We can see that all the simulation results are consistent with the theoretical values calculated by (12), which verifies the correctness of our theoretical model. Observe the trend that an increased information cost c can lead to a reduced probability of each UE sharing genuine information, which also matches the conditions derived in (12). In Figs. 6 and 7, we have fixed the information acquisition probability to $p_a = 0.8$. By contrast, we show the impact of the information acquisition probability on the specific value of the credibility equilibrium for the genuine information sharing cost c in Fig. 8. Explicitly, Fig. 8 shows that if all UEs actively acquire information with a higher probability, the probability of sharing genuine information would decrease, augmenting the selfish inclination.

C. Heterogeneous Model

We first simulated the heterogeneous model without relying on the credit mechanism, where all UEs are endowed with different initial mixed genuine/false information sharing strategies in the network. The reinforcement learning procedure characterized in Fig. 9 is implemented with the aid of 20 heterogeneous UEs, where the utility function is set to $\Gamma = 1$ and we have $\mu = 0.1$, while the discounting factor ϵ_i^t is simply set to $\frac{1}{n}$. In this heterogeneous simulation, we randomize the information sharing cost of each UE c_i within $[0.01, 0.09]$. Again, the learning curves are shown in Fig. 9, while the y-axis quantifies the Euclidean distance between the mixed strategies of two adjacent time instances, which is formulated as $\log\left(\sum_i ||\mathbf{s}_i^{t+1} - \mathbf{s}_i^t||_2\right)$. The three different lines correspond to different settings of the exploration levels ξ. Observe in the figure that the convergence performance of an aggressive exploration using $\xi = 0.1$ is worse than that of a slow exploration relying on $\xi = 10$.

We also show the converged credibility equilibrium associated with each UE's genuine information sharing cost in Fig. 10, where the UE index was set to be inversely proportional to the information sharing cost and the box chart represents each UE's information sharing strategies with respect to all the other 19 UEs. The exploration of the reinforcement learning model ensures that each UE has a non-zero probability of of sharing genuine information, instead of adopting a pure non-cooperative, selfish strategy. Generally, it can be seen in Fig. 10 that the UEs having a higher user index and hence a lower cooperation cost would share genuine information with a higher probability upon approaching the creditability equilibrium. Each box in Fig. 10 represents the $25\% - 75\%$ percentile of each UE's information sharing strategies with respect to other UEs, the star $*$ above each box represents each UE's strategy with respect to UE 20, while the star $*$ below each box represents each UE's strategy with respect to UE 1. Finally, the square $\square$ in the middle of the box represents each UE's average strategy and the horizontal line indicates the median. Apparently, when UE i shares information with UE j having a high index and a low information sharing cost (i.e. having a high probability of sharing genuine information), UE i also provides genuine information with a higher probability for UE j, which indicates an improved mutual trust. Observe in Fig. 10 that this phenomenon is dominated by each UE's own information sharing cost.

The heterogeneous model combined with our credit mechanism is also simulated by running Algorithm 1 over 20 UEs, where the reputation adjustment step size was configured according to $\frac{0.02}{t}$ with t being the time index. Fig. 11 shows the dynamics of all UEs' reputations during the learning and interaction process, which also characterizes the UEs' information sharing strategy. Although the UEs are initially configured to have different reputations below 0.5, i.e. to have a relatively low reputation, the final converged all "1" reputation results corroborate the high efficiency of our credit mechanism. As shown in Fig. 11, after the network becomes converged, some UEs might selfishly deviate from their good reputation at the index of 200, but they will promptly return to a good reputation again, which verifies the robustness of the proposed credit mechanism.

VII. Conclusions

In this paper, we studied a range of information credibility issues of cooperative networks, where both a homogeneous and a heterogeneous models were investigated. The credibility equilibria were derived under these two models, either in a closed-form expression based evolutionary game or using a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

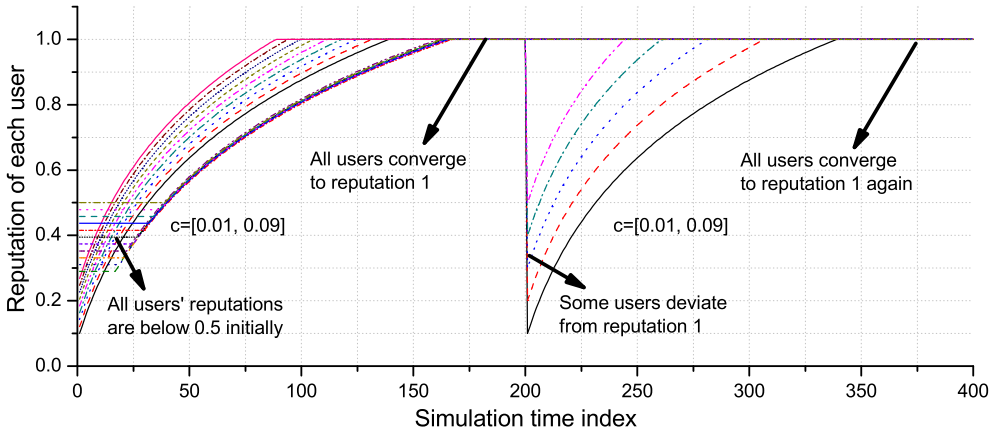

Fig. 11. Dynamics of the UEs' reputation.

reinforcement learning aided method. We also proposed a credit mechanism for enhancing the information credibility. Our simulation results showed that the cost of sharing genuine information had a grave impact on the credibility equilibrium. The proposed credit mechanism ensured that all UEs would share genuine information after a few rounds of interactions. Furthermore, 'defecting rogue' UEs became capable of rapidly recovering from sudden selfish deviations from cooperation. In a nutshell, we proposed a framework for the information credibility modelling of cooperative networks, which can be applied in diverse practical scenarios to be explored perhaps by you valued colleagues.

APPENDIX
PROOF OF THEOREM 1

Proof: In order to verify the conditions of those solutions being evolutionary credibility equilibria, we need to first analyze the characteristics of the first-order derivative of function $\Phi(P_{\mathbb{G}}) = \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}})$. According to (8) and (9), we have

$$\Phi(P_{\mathbb{G}}) = \sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n} (1-P_{\mathbb{G}})^{N-1-n} (1-p_a) \left[e^{-\mu n} - e^{-\mu(n+1)}\right] - c. \tag{44}$$

When $P_{\mathbb{G}} \in (0,1)$, the first-order derivative of $\Phi(P_{\mathbb{G}})$ can be calculated by

$$\Phi'(P_{\mathbb{G}}) = \sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] (1-p_a) \left[e^{-\mu n} - e^{-\mu(n+1)}\right]. \tag{45}$$

Let us define a threshold n_{th} satisfying that $n_{th} \leq (N-1)P_{\mathbb{G}}$ and $n_{th}+1 > (N-1)P_{\mathbb{G}}$. Then, we can re-write (45) by

$$\begin{aligned}\Phi'(P_{\mathbb{G}}) = & \sum_{n=0}^{n_{th}} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] (1-p_a) \left[e^{-\mu n} - e^{-\mu(n+1)}\right] \\ & + \sum_{n=n_{th}+1}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] (1-p_a) \left[e^{-\mu n} - e^{-\mu(n+1)}\right].\end{aligned} \tag{46}$$

Since the term $\left[e^{-\mu n} - e^{-\mu(n+1)}\right]$ is a decreasing function in terms of n, it holds that $\left[e^{-\mu n} - e^{-\mu(n+1)}\right] \geq \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right]$ when $n \leq n_{th}$ and $\left[e^{-\mu n} - e^{-\mu(n+1)}\right] < \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right]$ when $n > n_{th}$. In such a case, (46) satisfies that

$$\begin{aligned}\Phi'(P_{\mathbb{G}}) < & \sum_{n=0}^{n_{th}} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] (1-p_a) \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right] \\ & + \sum_{n=n_{th}+1}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] (1-p_a) \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right] \\ = & (1-p_a) \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right] \sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n-1} (1-P_{\mathbb{G}})^{N-2-n} [n-(N-1)P_{\mathbb{G}}] \\ = & (1-p_a) \left[e^{-\mu n_{th}} - e^{-\mu(n_{th}+1)}\right] \cdot \frac{\mathrm{d}\left[\sum_{n=0}^{N-1} \binom{N-1}{n} P_{\mathbb{G}}^{n} (1-P_{\mathbb{G}})^{N-1-n}\right]}{\mathrm{d}n} \\ = & 0.\end{aligned} \tag{47}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

When $P_{\mathbb{G}}=0$, we have

$$\Phi'(0)=\lim_{\epsilon\to 0}\frac{\Phi(\epsilon)-\Phi(0)}{\epsilon}=$$
$$\lim_{\epsilon\to 0}\frac{\sum_{n=0}^{N-1}\binom{N-1}{n}\epsilon^n(1-\epsilon)^{N-1-n}(1-p_a)\left[e^{-\mu n}-e^{-\mu(n+1)}\right]}{\epsilon}$$
$$-\frac{(1-p_a)\left(1-e^{-\mu}\right)}{\epsilon}=$$
$$\lim_{\epsilon\to 0}\frac{(1-\epsilon)^{N-1}(1-p_a)\left(1-e^{-\mu}\right)-(1-p_a)\left(1-e^{-\mu}\right)}{\epsilon}+$$
$$\lim_{\epsilon\to 0}\frac{(N-1)\epsilon(1-\epsilon)^{N-2}(1-p_a)\left(e^{-\mu}-e^{-2\mu}\right)}{\epsilon}$$
$$=(N-1)(1-p_a)\left(2e^{-\mu}-e^{-2\mu}-1\right)<0. \tag{48}$$

Similarly, when $P_{\mathbb{G}}=1$, we have

$$\Phi'(1)=\lim_{\epsilon\to 0}\frac{\Phi(1)-\Phi(1-\epsilon)}{\epsilon}= \tag{49}$$
$$\lim_{\epsilon\to 0}\frac{(1-p_a)\left[e^{-(N-1)\mu}-e^{-N\mu}\right]}{\epsilon}-$$
$$\frac{\sum_{n=0}^{N-1}\binom{N-1}{n}(1-\epsilon)^n\epsilon^{N-1-n}(1-p_a)\left[e^{-\mu n}-e^{-\mu(n+1)}\right]}{\epsilon}$$
$$=\lim_{\epsilon\to 0}\frac{\left[1-(1-\epsilon)^{N-1}\right](1-p_a)\left[e^{-(N-1)\mu}-e^{-N\mu}\right]}{\epsilon}-$$
$$\lim_{\epsilon\to 0}\frac{(N-1)\epsilon(1-\epsilon)^{N-2}(1-p_a)\left[e^{-(N-2)\mu}-e^{-(N-1)\mu}\right]}{\epsilon}$$
$$=(N-1)(1-p_a)\left[2e^{-(N-1)\mu}-e^{-N\mu}-e^{-(N-2)\mu}\right]<0,$$

where the last inequality is due to the decreasing function $\left[e^{-\mu n}-e^{-\mu(n+1)}\right]$ in terms of n. Therefore, combining (47), (48), and (49), we can see that $\Phi'(P_{\mathbb{G}})<0$ for all $P_{\mathbb{G}}\in[0,1]$. Based on this characteristic, let us verify the conditions of the three solutions $P_{\mathbb{G}}^{*1}=0$, $P_{\mathbb{G}}^{*2}=1$ and $P_{\mathbb{G}}^{*3}=\left(1-\frac{c}{(1-p_a)(1-e^{-\mu})}\right)^{\frac{1}{N-1}}\Big/\left(1-e^{-\mu}\right)$ being evolutionary credibility equilibria as follows.

- Solution 1: $P_{\mathbb{G}}^{*1}=0$

When a UE shares genuine information with probability $P_{\mathbb{G}}$, while all other UEs share false information with probability 1, the average utility according to (10) is

$$\bar{U}(P_{\mathbb{G}},0)=\bar{U}_{\mathbb{F}}(0)+P_{\mathbb{G}}\left[\bar{U}_{\mathbb{G}}(0)-\bar{U}_{\mathbb{F}}(0)\right], \tag{50}$$

where

$$\bar{U}_{\mathbb{G}}(0)=\Gamma-e^{-\mu}-c, \tag{51}$$
$$\bar{U}_{\mathbb{F}}(0)=\Gamma+p_a(1-e^{-\mu})-1, \tag{52}$$

according to (8) and (9), respectively. If $\frac{c}{1-p_a}<1-e^{-\mu}$, then it holds that $\bar{U}_{\mathbb{G}}(0)>\bar{U}_{\mathbb{F}}(0)$ and thus all UEs would transit to sharing genuine information for the sake of obtaining higher utility $\bar{U}_{\mathbb{G}}(0)$. On the other hand, if $\frac{c}{1-p_a}>1-e^{-\mu}$, then it holds that $\bar{U}_{\mathbb{G}}(0)<\bar{U}_{\mathbb{F}}(0)$ and thus all UEs would remain current strategy, i.e. keeping sharing false information for the sake of obtaining higher utility $\bar{U}_{\mathbb{F}}(0)$. While for the boundary $\frac{c}{1-p_a}=1-e^{-\mu}$, i.e. $\bar{U}_{\mathbb{G}}(0)=\bar{U}_{\mathbb{F}}(0)$, it holds that $\bar{U}(P_{\mathbb{G}},0)=\bar{U}_{\mathbb{F}}(0)=\bar{U}(0,0)$, which means $P_{\mathbb{G}}^{*1}=0$ satisfies the NE property in *Definition 1*. Let us further check the stability condition. Since $\Phi(0)=\bar{U}_{\mathbb{G}}(0)-U_{\mathbb{F}}(0)=0$ and $\Phi'(P_{\mathbb{G}})<0\ \forall P_{\mathbb{G}}\in[0,1]$, we have $\Phi(P_{\mathbb{G}})<0\ \forall P_{\mathbb{G}}\in[0,1]$. Thus, it holds that

$$\begin{aligned}\bar{U}(P_{\mathbb{G}},P_{\mathbb{G}})&=\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})+P_{\mathbb{G}}\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})-\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right]\\&<\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})=\bar{U}(0,P_{\mathbb{G}}),\end{aligned} \tag{53}$$

which implies $P_{\mathbb{G}}^{*1}=0$ also satisfies the stability property in *Definition 1*. Therefore, the condition of $P_{\mathbb{G}}^{*1}=0$ being an evolutionary credibility equilibrium is $\frac{c}{1-p_a}\geq 1-e^{-\mu}$.

- Solution 2: $P_{\mathbb{G}}^{*2}=1$

When a UE shares genuine information with probability $P_{\mathbb{G}}$, while all other UEs share genuine information with probability 1, the average utility according to (10) is

$$\bar{U}(P_{\mathbb{G}},1)=\bar{U}_{\mathbb{F}}(1)+P_{\mathbb{G}}\left[\bar{U}_{\mathbb{G}}(1)-\bar{U}_{\mathbb{F}}(1)\right], \tag{54}$$

where

$$\bar{U}_{\mathbb{G}}(1)=\Gamma-e^{-\mu N}-c, \tag{55}$$
$$\bar{U}_{\mathbb{F}}(1)=\Gamma-e^{-\mu(N-1)}+p_a\left[e^{-\mu(N-1)}-e^{-\mu N}\right], \tag{56}$$

according to (8) and (9), respectively. If $\frac{c}{1-p_a}>e^{-\mu(N-1)}-e^{-\mu N}$, then it holds that $\bar{U}_{\mathbb{G}}(1)<\bar{U}_{\mathbb{F}}(1)$ and thus all UEs would transit to sharing false information for the sake of obtaining higher utility $\bar{U}_{\mathbb{F}}(1)$. On the other hand, if $\frac{c}{1-p_a}<e^{-\mu(N-1)}-e^{-\mu N}$, then it holds that $\bar{U}_{\mathbb{G}}(1)>\bar{U}_{\mathbb{F}}(1)$ and thus all UEs would remain current strategy, i.e. keeping sharing genuine information for the sake of obtaining higher utility $\bar{U}_{\mathbb{G}}(1)$. While for the boundary $\frac{c}{1-p_a}=e^{-\mu(N-1)}-e^{-\mu N}$, i.e. $\bar{U}_{\mathbb{G}}(1)=\bar{U}_{\mathbb{F}}(1)$, it holds that $\bar{U}(P_{\mathbb{G}},1)=\bar{U}_{\mathbb{F}}(1)=\bar{U}(1,1)$, which means $P_{\mathbb{G}}^{*2}=1$ satisfies the NE property in *Definition 1*. Let us further check the stability condition. Since $\Phi(1)=\bar{U}_{\mathbb{G}}(1)-U_{\mathbb{F}}(1)=0$ and $\Phi'(P_{\mathbb{G}})<0\ \forall P_{\mathbb{G}}\in[0,1]$, we have $\Phi(P_{\mathbb{G}})>0\ \forall P_{\mathbb{G}}\in[0,1]$. Thus, it holds that

$$\begin{aligned}\bar{U}(P_{\mathbb{G}},P_{\mathbb{G}})&=\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})+P_{\mathbb{G}}\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})-\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right]\\&<\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})+1\cdot\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})-\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right]=\bar{U}(1,P_{\mathbb{G}}),\end{aligned} \tag{57}$$

which implies that $P_{\mathbb{G}}^{*2}=1$ also satisfies the stability property in *Definition 1*. Therefore, the condition of $P_{\mathbb{G}}^{*2}=1$ being an evolutionary credibility equilibrium is $\frac{c}{1-p_a}\leq e^{-\mu(N-1)}-e^{-\mu N}$.

- $P_{\mathbb{G}}^{*3}=\left(1-\frac{c}{(1-p_a)(1-e^{-\mu})}\right)^{\frac{1}{N-1}}\Big/\left(1-e^{-\mu}\right)$

When UE i shares genuine information with probability $P_{\mathbb{G}}$, while all other UEs share genuine information with probability $P_{\mathbb{G}}^{*3}$, the average utility of UE i can be calculated by

$$\bar{U}_i(P_{\mathbb{G}},P_{\mathbb{G}}^{*3})=P_{\mathbb{G}}\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{*3})+(1-P_{\mathbb{G}})\bar{U}_{\mathbb{F}}(P_{\mathbb{G}}^{*3}). \tag{58}$$

Since $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{*3})-\bar{U}_{\mathbb{F}}(P_{\mathbb{G}}^{*3})=0$ according to (11), we can re-write (58) as

$$\bar{U}_i(P_{\mathbb{G}},P_{\mathbb{G}}^{*3})=\bar{U}_{\mathbb{F}}(P_{\mathbb{G}}^{*3})=\bar{U}_i(P_{\mathbb{G}}^{*3},P_{\mathbb{G}}^{*3}), \tag{59}$$

which implies that $P_{\mathbb{G}}^{*3}$ satisfies the NE property in *Definition 1*. Let us further check the stability condition. According to (10), we have

$$\bar{U}_i(P_{\mathbb{G}},P_{\mathbb{G}})=\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})+P_{\mathbb{G}}\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}})-\bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right], \tag{60}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

and

$$\bar{U}_i(P_{\mathbb{G}}^{*3}, P_{\mathbb{G}}) = \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) + P_{\mathbb{G}}^{*3}\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right]. \quad (61)$$

By combining (60) and (61), we have

$$\bar{U}_i(P_{\mathbb{G}}^{*3}, P_{\mathbb{G}}) - \bar{U}_i(P_{\mathbb{G}}, P_{\mathbb{G}}) = \left(P_{\mathbb{G}}^{*3} - P_{\mathbb{G}}\right)\left[\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}})\right]. \quad (62)$$

Again, based on the characteristic of $\Phi'(P_{\mathbb{G}}) = \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) < 0\ \forall P_{\mathbb{G}} \in [0,1]$, we have

1) if $P_{\mathbb{G}} < P_{\mathbb{G}}^{*3}$, then $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) > \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{*3}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}^{*3}) = 0$;
2) if $P_{\mathbb{G}} > P_{\mathbb{G}}^{*3}$, then $\bar{U}_{\mathbb{G}}(P_{\mathbb{G}}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}) < \bar{U}_{\mathbb{G}}(P_{\mathbb{G}}^{*3}) - \bar{U}_{\mathbb{F}}(P_{\mathbb{G}}^{*3}) = 0$.

Therefore, it holds that $\bar{U}_i(P_{\mathbb{G}}^{*3}, P_{\mathbb{G}}) < \bar{U}_i(P_{\mathbb{G}}, P_{\mathbb{G}})\ \forall P_{\mathbb{G}} \neq P_{\mathbb{G}}^{*3}$, which indicates that $P_{\mathbb{G}}^{*3}$ also satisfies the stability property in *Definition 1*. This completes the proof of *Theorem 1*.

APPENDIX
PROOF OF THEOREM 2

Proof: Let us first introduce some definitions to further reveal the physical meaning of (28). Firstly, we denote $\Theta_i = \{\mathbb{G}, \mathbb{F}\}$ as the set of UE i's pure strategy and denote $\Theta = \prod_{i\in N}\Theta_i$ as the space of all UEs' pure strategy profile. Meanwhile, let us denote $\Lambda_i = [0,1]$ as the set of UE i's mixed strategy, i.e. the probability of sharing genuine information, and denote $\Lambda = \prod_{i\in N}\Lambda_i$ as the space of all UEs' mixed strategy profile. Then, we consider the space of perception $\Xi = \prod_{i\in N}\mathbb{R}^{\Theta_i}$ and a mapping $\chi : \Xi \to \Lambda$ from all UEs' perceptions to the mixed strategy profile of all UEs, $\chi(\mathbf{P}) = [\phi_{ij}(\mathbf{P}_i)]_{i\in N}$, where $\phi_{ij} : \mathbb{R}^{\Theta_i} \to \Lambda_{ij}$ is a continuous mapping from UE i's space of perceptions to its space of mixed strategies to UE j, which is just the operation defined in (24). Moreover, we introduce a self mapping $\Pi : \Xi \to \Xi$ as a function of the perception

$$\Pi(\mathbf{P}) = F\left[\chi(\mathbf{P})\right], \quad (63)$$

where $F : \Lambda \to \Xi$ is a mapping from mixed strategy profile to the perception, i.e. $F(\mathbf{s}) = [F_{ij}(\mathbf{s})]_{i\in N}$ and $F_{ij}(\mathbf{s}) = [F_{ij\mathbb{G}}(\mathbf{s}), F_{ij\mathbb{F}}(\mathbf{s})]_{i\in N}$ with $\mathbf{s}$ being the profile of all UEs' mixed strategies. As a matter of fact, $F_i(\mathbf{s})$ is UE i's average utility when it adopts a pure strategy, i.e.

$$F_{ij\mathbb{G}}(\mathbf{s}) = \bar{U}_{ij}(\mathbb{G}, \mathbf{s}^{-i}), \quad F_{ij\mathbb{F}}(\mathbf{s}) = \bar{U}_{ij}(\mathbb{F}, \mathbf{s}^{-i}), \quad (64)$$

where $\mathbf{s}^{-i}$ denotes the strategy profile except UE i. Based on those definitions, we can re-write the expectation term in (28) as follows

$$\begin{aligned}\mathbb{E}(q_{ij\mathbb{G}}|\mathbf{P}) &= s_{ij}\bar{U}_{ij}(\mathbb{G}, \mathbf{s}^{-i}) + (1 - s_{ij})p_{ij\mathbb{G}} \quad (65)\\ &= \phi_{ij\mathbb{G}}(\mathbf{P})\Pi_{ij\mathbb{G}}(\mathbf{P}) + \left[1 - \phi_{ij\mathbb{G}}(\mathbf{P})\right]p_{ij\mathbb{G}}.\end{aligned}$$

By inserting (65) into (28), we have

$$\dot{p}_{ij\mathbb{G}} = \phi_{ij\mathbb{G}}(\mathbf{P})\left[\Pi_{ij\mathbb{G}}(\mathbf{P}) - p_{ij\mathbb{G}}\right], \quad (66)$$

which is quite similar to the replicator dynamics in (7) of the homogeneous scenario. Similar to *Theorem 1*, the convergence of the reinforcement learning process should accompany with the dynamics $\dot{p}_{ij\mathbb{G}}$ in (66) approaching 0, i.e. $\Pi_{ij\mathbb{G}}(\mathbf{P}) = p_{ij\mathbb{G}}$. In such a case, the condition of the mapping $\Pi_{ij\mathbb{G}}(\cdot)$ having a fix point becomes the convergence condition of the proposed reinforcement learning scheme.

Given two arbitrary perceptions $\mathbf{P}_1$ and $\mathbf{P}_2$ of, the difference of $\Pi_{ij\mathbb{G}}(\mathbf{P}_1) - \Pi_{ij\mathbb{G}}(\mathbf{P}_2)$ can be calculated as follows:

$$\begin{aligned}|\Pi_{ij\mathbb{G}}(\mathbf{P}_1) - \Pi_{ij\mathbb{G}}(\mathbf{P}_2)| =& \Big|\chi_i(\mathbf{P}_1)\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right] -\\ &\chi_i(\mathbf{P}_2)\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right] +\\ &\chi_j(\mathbf{P}_1)\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right] -\\ &\chi_j(\mathbf{P}_2)\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right]\Big|\\ \le& \Big|\left[\chi_i(\mathbf{P}_1) - \chi_i(\mathbf{P}_2)\right]\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right]\Big| +\\ &\Big|\left[\chi_j(\mathbf{P}_1) - \chi_j(\mathbf{P}_2)\right]\left[U_{ij}(\mathbb{G},\mathbb{G}) - U_{ij}(\mathbb{G},\mathbb{F})\right]\Big|\\ =& (e^{-\mu} - e^{-2\mu} + g)\Big[\left|\chi_i(\mathbf{P}_1) - \chi_i(\mathbf{P}_2)\right| +\\ &\left|\chi_j(\mathbf{P}_1) - \chi_j(\mathbf{P}_2)\right|\Big]. \quad (67)\end{aligned}$$

According to the Boltzmann update rule in (24), we can derive the difference of $\chi_i(\mathbf{P}_1) - \chi_i(\mathbf{P}_2)$ as follows

$$\begin{aligned}|\chi_i(\mathbf{P}_1) - \chi_i(\mathbf{P}_2)| =& \left|\frac{e^{\xi p_{1i\mathbb{G}}}}{e^{\xi p_{1i\mathbb{G}}} + e^{\xi p_{1i\mathbb{F}}}} - \frac{e^{\xi p_{2i\mathbb{G}}}}{e^{\xi p_{2i\mathbb{G}}} + e^{\xi p_{2i\mathbb{F}}}}\right|\\ =& \Bigg|\xi\frac{e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}}} + e^{\xi\overline{p}_{1i\mathbb{F}}}\right)^2}(p_{1i\mathbb{G}} - p_{2i\mathbb{G}}) -\\ &\xi\frac{e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}}} + e^{\xi\overline{p}_{1i\mathbb{F}}}\right)^2}(p_{1i\mathbb{F}} - p_{2i\mathbb{F}})\Bigg|\\ \le& \xi\frac{e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}}} + e^{\xi\overline{p}_{1i\mathbb{F}}}\right)^2}\left|p_{1i\mathbb{G}} - p_{2i\mathbb{G}}\right| +\\ &\xi\frac{e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}} + e^{\xi\overline{p}_{1i\mathbb{F}}}}\right)^2}\left|p_{1i\mathbb{F}} - p_{2i\mathbb{F}}\right|\\ \le& \xi\frac{2e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}} + e^{\xi\overline{p}_{1i\mathbb{F}}}}\right)^2}||\mathbf{P}_{1i} - \mathbf{P}_{2i}||_\infty\\ \le& \xi||\mathbf{P}_{1i} - \mathbf{P}_{2i}||_\infty, \quad (68)\end{aligned}$$

where the second equality is according to the mean value theorem and the last step is due to $\frac{2e^{\xi(\overline{p}_{1i\mathbb{G}} + \overline{p}_{1i\mathbb{F}})}}{\left(e^{\xi\overline{p}_{1i\mathbb{G}} + e^{\xi\overline{p}_{1i\mathbb{F}}}}\right)^2} \le 1$. By inserting (68) into (67), we have

$$|\Pi_{ij\mathbb{G}}(\mathbf{P}_1) - \Pi_{ij\mathbb{G}}(\mathbf{P}_2)| \le 2\xi(e^{-\mu} - e^{-2\mu} + g)||\mathbf{P}_1 - \mathbf{P}_2||_\infty. \quad (69)$$

We can see that if $\xi \le \frac{1}{2(e^{-\mu} - e^{-2\mu} + g)}$, the mapping Π forms a maximum-norm contraction, which can guarantee the existence of a unique fixed point. This completes the proof of *Theorem 2*.

REFERENCES

[1] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, "Emergence of cooperation and evolutionary stability in finite populations," *Nature*, vol. 428, no. 6983, pp. 646–650, Apr. 2004.
[2] I. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, Apr. 2008.
[3] J. Ahny, Y. Wang, B. Yu, F. Bai, and B. Krishnamachari, "Risa: Distributed road information sharing architecture," in *Proc. IEEE INFOCOM*, Orlando, FL, Mar. 2012, pp. 1494–1502.

[4] J. Zhang, R. Zhang, G. Li, and L. Hanzo, "Remote coalition network elements for base station cooperation aided multicell processing," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1406–1415, Mar. 2012.

[5] X. Li, R. Zhang, and L. Hanzo, "Cooperative load balancing in hybrid visible light communications and wifi," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1319–1329, Apr. 2015.

[6] F. Li, B. Luo, P. Liu, D. Lee, and C.-H. Chu, "Enforcing secure and privacy-preserving information brokering in distributed information sharing," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 888–900, Jun. 2013.

[7] S. Song, K. Hwang, R. Zhou, and Y.-K. Kwok, "Trusted p2p transactions with fuzzy reputation aggregation," *IEEE Internet Comput.*, vol. 9, no. 6, pp. 24–34, Dec. 2005.

[8] L. Cai and R. Rojas-Cessa, "Bounding virus proliferation in p2p networks with a diverse-parameter trust management scheme," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 812–814, Oct. 2009.

[9] J. Weng, Z. Shen, C. Miao, A. G. E. Soong, and C. Leung, "Credibility: How agents can handle unfair third-party testimonies in computational trust models," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 9, pp. 1286–1298, Sep. 2010.

[10] B. Liu, P. Terlecky, A. Bar-Noy, R. Govindan, M. J. Neely, and D. Rawitz, "Optimizing information credibility in social swarming applications," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 6, pp. 1147–1158, Jun. 2012.

[11] L. Canzian, Y. Xiao, W. Zame, M. Zorzi, and M. van der Schaar, "Intervention with private information, imperfect monitoring and costly communication," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3192–3205, Aug. 2013.

[12] Y. Sun, W. Yu, Z. Han, and K. Liu, "Information theoretic framework of trust modeling and evaluation for ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 305–317, Feb. 2006.

[13] A. Gopinathan, Z. Li, and B. Li, "Group strategyproof multicast in wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 5, pp. 708–715, May 2011.

[14] Y. Zou, X. Wang, W. Shen, and L. Hanzo, "Security versus reliability analysis of opportunistic relaying," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2653–2661, Jun. 2013.

[15] Z. Jin, S. Anand, and K. Subbalakshmi, "Impact of primary user emulation attacks on dynamic spectrum access networks," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2635–2643, Sep. 2012.

[16] R. Vaidyanathaswami and A. Thangaraj, "Robustness of physical layer security primitives against attacks on pseudorandom generators," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1070–1079, Mar. 2014.

[17] G. D. Tormo, F. G. Marmol, J. Girao, and G. M. Perez, "Identity management–in privacy we trust: Bridging the trust gap in ehealth environments," *IEEE Security Privacy*, vol. 11, no. 6, pp. 34–41, Dec. 2013.

[18] M. Nitti, R. Girau, and L. Atzori, "Trustworthiness management in the social internet of things," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 5, pp. 1253–1266, May 2014.

[19] Y. Chen and K. J. R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Transactions on Communications*, vol. 59, no. 1, pp. 159–168, January 2011.

[20] B. Niu, H. V. Zhao, and H. Jiang, "A cooperation stimulation strategy in wireless multicast networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2355–2369, May 2011.

[21] Y. Chen and K. J. R. Liu, "Understanding microeconomic behaviors in social networking: An engineering view," *IEEE Signal Processing Magazine*, vol. 29, no. 2, pp. 53–64, March 2012.

[22] H. Chen, W. Lou, Z. Wang, and Q. Wang, "A secure credit-based incentive mechanism for message forwarding in noncooperative dtns," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6377–6388, Aug 2016.

[23] H. y. Lee and Y. b. Lin, "Credit pre-reservation mechanism for umts prepaid service," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1867–1873, June 2010.

[24] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[25] N. Alam, A. T. Balaei, and A. G. Dempster, "A dsrc doppler-based cooperative positioning enhancement for vehicular networks with gps availability," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4462–4470, Sep. 2011.

[26] M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, "Collaborative filtering recommender systems," *Foundations and Trends in Human-Computer Interaction*, vol. 4, no. 2, pp. 81–173, Feb. 2011.

[27] M. Bouakiz and M. J. Sobel, "Inventory control with an exponential utility criterion," *Operations Research*, vol. 40, no. 3, pp. 603–608, Jun. 1992.

[28] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. MIT Press, 2003.

[29] C. Jiang, Y. Chen, and K. J. R. Liu, "Graphical evolutionary game for information diffusion over social networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 524–536, 2014.

[30] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 2008–2017, May 2009.

[31] C. Jiang, Y. Chen, Y. Gao, and K. Liu, "Joint spectrum sensing and access evolutionary game in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2470–2483, May 2013.

[32] X. Chen and J. Huang, "Evolutionarily stable spectrum access," *IEEE Trans. Mobile Compt.*, vol. 12, no. 7, pp. 1281–1293, Jul. 2013.

[33] C. Jiang, Y. Chen, and K. J. R. Liu, "Distributed adaptive networks: a graphical evolutionary game theoretic view," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5675–5688, 2013.

[34] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, "Cooperative peer-to-peer streaming: An evolutionary game-theoretic approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 10, pp. 1346–1357, Oct 2010.

[35] C. Jiang, Y. Chen, and K. J. R. Liu, "Evolutionary dynamics of information diffusion over social networks," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4573–4586, 2014.

[36] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT Press, 1998.

[37] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, "Self-organization in small cell networks: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3202–3212, Jul. 2013.

[38] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Syst. Mag.*, vol. 32, no. 6, pp. 76–105, Dec. 2012.

[39] T. Matsui, T. Goto, K. Izumi, and Y. Chen, *Recent Advances in Reinforcement Learning*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, vol. 7188, ch. Compound Reinforcement Learning: Theory and an Application to Finance, pp. 321–332.

[40] R. M. Jones, L. H. Somerville, J. Li, E. J. Ruberry, V. Libby, G. Glover, H. U. Voss, D. J. Ballon, and B. J. Casey, "Behavioral and neural properties of social reinforcement learning," *The Journal of Neuroscience*, vol. 31, no. 37, pp. 13 039–13 045, Sep. 2011.

[41] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2578–2590, Dec 2015.

[42] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, "A mobile offloading game against smart attacks," *IEEE Access*, vol. 4, pp. 2281–2291, 2016.

[43] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "Phy-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2016.

[44] K. Goldberg, "Jester joke recommender system," [Online]: http://grouplens.org/datasets/movielens/1m/.

[45] K. Bache and M. Lichman, "Uci machine learning repository," [Online]: http://archive.ics.uci.edu/ml.

[46] F. Kohlmayer, F. Prasser, C. Eckert, A. Kemper, and K. Kuhn, "Flash: Efficient, stable and optimal k-anonymity," in *2012 International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2012 International Confernece on Social Computing (SocialCom)*, Amsterdam, Netherland, 2012, pp. 708–717.

[47] B. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing: A survey of recent developments," *ACM Computing Surveys*, vol. 42, no. 4, pp. 14.1–14.53, Jun. 2010.

[48] V. S. Iyengar, "Transforming data to satisfy privacy constraints," in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, USA, 2002, pp. 279–288.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2017.2659078, IEEE Journal on Selected Areas in Communications

Chunxiao Jiang (S'09-M'13-SM'15) received the B.S. in information engineering from Beihang University in Jun. 2008 and the Ph.D. in electronic engineering from Tsinghua University in Jan. 2013, both with the highest honors. From Feb. 2013 - Jun. 2016, Dr. Jiang was a Postdoc in the Department of Electronic Engineering Tsinghua University, during which he visited University of Maryland College Park and University of Southampton. He is a recipient of the Best Paper Award from IEEE Globecom 2013 and the Best Student Paper Award from IEEE GlobalSIP 2015. Since 2015, Dr. Jiang became a IEEE Senior Member.

Linling Kuang (S'01CM'06) received the B.S. and M.S. degrees from the National University of Defense Technology, Changsha, China, in 1995 and 1998, respectively, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004. Since 2007, she has been with the Tsinghua Space Center, Tsinghua University. Her research interests include wireless broadband communications, signal processing, and satellite communication. Dr. Kuang is a member of the IEEE Communications Society.

Zhu Han (S'01-M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Currently, Dr. Han is an IEEE Communications Society Distinguished Lecturer.

Yong Ren (SM'15) 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 in Tsinghua University. He holds 12 patents, and has authored or co-authored more than 100 technical papers in the behavior of computer network, P2P network and cognitive networks. He has serves as a reviewer of IEICE Transactions on Communications, Digital Signal Processing, Chinese Physics Letters, Chinese Journal of Electronics, Chinese Journal of Computer Science and Technology, Chinese Journal of Aeronautics and so on. 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.

Lajos Hanzo FREng, FIEEE, FIET, Fellow of EURASIP, DSc received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 111 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1600+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008-2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk. Lajos has 26 000+ citations and an H-index of 62.