Information Reliability in Complex Multitask Networks

Sadaf Monajemi[a,*], Saeid Sanei[b], Sim-Heng Ong[c]

[a]*NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore*
[b]*Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK*
[c]*Electrical and Computer Engineering, Faculty of Engineering, National University of Singapore, Singapore*

Abstract

The emergence of distributed and complex networks has altered the field of information and data processing in the past few years. In distributed networks, the connected neighboring nodes can cooperate and share information with each other in order to solve particular tasks. However, in many applications the agents might be reluctant to share their true data with all their neighbors due to privacy and security constraints. In this paper, we study the performance of multitask distributed networks where sharing genuine information is subject to a cost. We formulate an information credibility model which results in the probability of sharing genuine information at each time instant according to the cost. Each agent then shares its true information with only a subset of its neighbors while sending fabricated data to the rest according to this probability. This behavior can affect the performance of the whole network in an adverse manner especially in cases where the cost is high. To overcome this problem, we propose an adaptive reputation protocol which enables the agents to evaluate the behavior of their neighbors over time and select the most reputable subset of neighbors to share genuine information with. We provide an extensive simulation-based analysis to compare the performance of the proposed method with several other distributed learning strategies. The results show that the proposed method outperforms the other learning strategies and enables the network to have a superior performance especially when the cost of sharing genuine information is high.

*Corresponding Author:
Email address: sadaf.monajemi@u.nus.edu (Sadaf Monajemi)

Keywords: Adaptive distributed networks, Information and communication security, Reputation protocol, Information credibility, Diffusion adaptation

1. Introduction

Decentralized systems have attracted much interest in several situations as information processing is done in a distributed and collaborative manner. Following these strategies, a set of spatially distributed agents who are linked to each other form an adaptive network. These nodes or agents of the network have local interactions with other agents which enable them to have cooperation. The performance of these self-organized networks depends on the learning abilities and the localized cooperation of the interconnected nodes. In these types of networks, each agent can communicate and share information with its neighboring nodes. As a result of this cooperation and information sharing, the agents can solve particular tasks or optimization problems (such as estimating unknown parameters, tracking objects, etc). Several strategies have been proposed for distributed information processing over networks, such as incremental [5, 37, 24, 31, 6], consensus [16, 8, 33, 32] and diffusion strategies [14, 13, 29, 25, 35, 27]. In the incremental strategy, a Hamiltonian cycle has to be determined across the nodes of the network, which is generally an NP-hard task [22]. Therefore, topology changes in the network over time presents a considerable obstacle for incremental methods. On the contrary, it has been shown that among these strategies, diffusion algorithm is robust, scalable, and capable of real-time adaptation and learning. Diffusion strategies have also superior performance and stability compared to consensus methods in data processing applications [35, 38].

The diffusion strategy and its performance has been studied extensively in several scenarios [14, 35, 28]. In most of the prior studies, it is assumed that the agents are genuine, trustworthy and obey certain distributed information sharing protocols which may be a strong limitation in real-world applications [14, 38, 11, 12]. Although sharing genuine information and being cooperative ~~is an essential element~~ for the efficiency of these networks, there are many types of cooperative networks (such as online social networks, websites, and social forums) where some of the agents might disobey the protocols, stop cooperating, or feed others with malicious and misleading information [39, 40, 41]. One of the reasons causing this behavior is

the fact that sharing genuine information can be sensitive due to privacy preservation in several situations [23, 10, 43]. In some other scenarios, selfish agents might prefer to disobey the rules and share misleading information to minimize their own costs as sharing genuine information can be expensive [19, 1, 20]. Moreover, some agents might share misleading information with others in order to obtain an unfair advantage or bias due to their own superior performance. Therefore, it is crucially important to consider the freedom of agents for sharing genuine information in order to analyze the behavior and performance of these networks.

Several research articles in this area ~~have been~~ dedicated to incentive mechanisms that encourage the agents to cooperate with their neighbors. Some of these mechanisms are based on pricing strategies where payments are used to reward or punish the agents for their behavior [26, 4]. These mechanisms might be adequate for some settings, but are not appropriate for several cases where the information is free such as online social networks. Moreover, these systems usually require complex monitoring and accounting infrastructures that ~~results~~ in significant computation costs. In other studies, differential services are considered to reward and punish the agents based on their actions. These services can be provided by a network operator in case of centralized systems [17, 41]. However, as there is no central operator in decentralized systems, differential services are provided by the other agents based on their interactions. In reputation mechanisms, a reputation score is assigned to an agent based on its past behavior with the other agents. However, in most of the studies on reputation mechanisms the focus is mainly on practical implementation details [3, 34]. In [39], the authors studied a network where agents provide different services to their neighbors and designed incentive-compatible rating systems. In another work [40], the authors studied the case where the agents are self-interested and try to minimize their own cost and estimation error over a ~~single task~~ network. In this work, the agents are randomly paired with each other and each agent can share information with only one agent at a time. In [41], a centralized rating protocol is formulated where all the agents have the incentive to follow the recommended strategy. In [19], the authors studied energy expenditure of agents in a cooperative network. They proposed a game-theoretic approach to help the agents decide about their activation based on the trade-off between their contribution and energy expenditure over a single task network.

This work differs from the existing methods in the literature from several aspects. First, we study multitask networks where there are several con-

nected clusters of nodes with different objectives. To make it more realistic, we remove any presumption about prior clustering information. Specifically, the agents have no former knowledge regarding the cluster that they or their neighbors belong to. Moreover, we do not assume that the objectives of the clusters are necessarily related. Second, we study the case where sharing genuine information is subjected to a cost and each node can decide whether to share genuine or misleading information. To consider this problem in a more generalized way, we allow the nodes to make this decision in a pairwise manner rather than holistic approaches where genuine or false information is sent to all the neighbors. This means that each node can share genuine information with a subset of its neighbors while sending misleading information to the remaining neighbors at each time instant. Third, we define utility functions for sharing genuine and misleading information to obtain a credibility model for the cooperative network. Credibility equilibrium can be found using the model which determines the probability of sharing genuine information for each agent. It is obvious that a low probability of genuine information sharing results in a degraded global benefit for the cooperative network. Lastly, we propose a reputation approach which allows the agents to evaluate the importance of their neighbors for their own estimation task. This is the first time that a reputation protocol has been incorporated in the multitask diffusion strategy. Considering the spontaneous behavior of the nodes and using the reputation scores, each node can select the subset of neighbors to share genuine information with according to the credibility equilibrium. With the help of this dynamic and adaptive protocol, each node would be able to select its most important and trustworthy neighbors to share information with while taking into account its own privacy and cost budgets. Table 1 highlights the differences of our method from the existing works.

The rest of this paper is organized as follows. In Section 2 we introduce our system model and the multitask diffusion adaptation strategy for information processing over the network. In Section 3, we introduce the utility functions and derive the credibility equilibrium. We propose the adaptive reputation protocol for sharing information in Section 4 and provide the simulation results in Section 5. The paper is concluded in Section 6.

2. System Model

First we provide a summary of some of the main symbols and notations that are used in this article. Other symbols are defined in the context where

	[14, 38]	[11, 12]	[39, 40]	[41]	[19]	This Work
Objective	Single-task	Multitask	Incentives to Coop	Incentives to Coop	Activation Control	Info. Reliability
Network Topology	Arbitrary	Arbitrary	Arbitrary, One-to-one	Random Matching	Arbitrary	Arbitrary
Multitask Network	No	Yes	No	No	No	Yes
Clustering Info.	N/A	Known	N/A	N/A	N/A	Unknown
Reliability Model	No	No	No	No	No	Yes
Info. Exchange	Costless	Costless	Costly	Costly	Costly	Costly
Utility Depends	N/A	N/A	Own and Others	Own and Others	Value and Cost	Own and Others
Proposed Protocol	N/A	N/A	Distr.	Centr.	Distr.	Distr.

Table 1: Comparison with existing works.

they are used:

$\mathcal{N}_k$	Neighbors of node k including node k
$d_k(i)$	Scalar measurement of node k at time i
$\boldsymbol{x}_k(i)$	Regression vector of node k at time i
$\boldsymbol{\omega}_k^o$	Optimum parameter vector of node k at time i
$\mathcal{C}_q$	Set of nodes that belong to cluster q
$\mathcal{C}(k)$	The cluster that node k belongs to
$a_{\ell k}(i)$	Weight assigned by node k to the information of node ℓ
$\mathcal{N}_{\mathbb{G}}^{\ell}(i)$	The subset that node ℓ shares its genuine information with
c_k	Cost of sharing genuine information
$p_{k,\mathbb{G}}^i$	Probability of sharing genuine information for agent k
$R_{\ell k}(i)$	Reputation score assigned to node ℓ by node k

2.1. Network Model

Consider a connected network of N nodes as shown in Figure 1. As can be seen in this figure, each node k is connected to a number of neighboring nodes represented by $\mathcal{N}_k$. Each agent k of the network has access to the scalar measurements $d_k(i)$ and a $M \times 1$ regression vector $\boldsymbol{x}_k(i)$ with covariance matrix $\boldsymbol{R}_{x,k} = \mathbb{E}\boldsymbol{x}_k(i)\boldsymbol{x}_k^*(i) > 0$ at every time instant i and M represents the dimension of the problem at hand. It is assumed that each node is interested to estimate a $M \times 1$ unknown parameter vector $\boldsymbol{\omega}_k^o$ that is related to the data $\{d_k(i), \boldsymbol{x}_k(i)\}$ via a linear regression model:

$$d_k(i) = \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_k^o + n_k(i), \tag{1}$$

where $n_k(i)$ is the measurement noise of node k at time instant i. To better understand this linear regression model, we present a physical example from [36], where a network of agents are spread over a geographical area observing realizations of an auto-regressive (AR) random process $d_k(i)$ of order M. The AR process observed by agent k satisfies the model:

$$d_k(i) = \sum_{m=1}^{M} \alpha_m d_k(i-m) + n_k(i), \quad k = 1, 2, \cdots N \tag{2}$$

where the scalars $\{\alpha_m\}$ are the model parameters that the agents are interested to identify, and $n_k(i)$ is the additive noise. The parameters $\{\alpha_m\}$ can

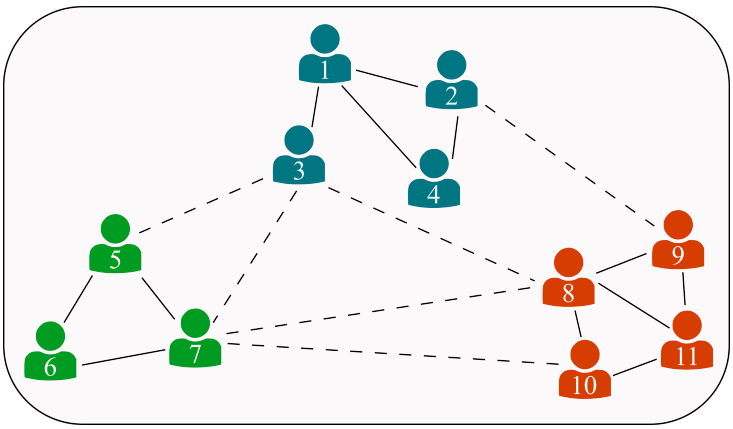

Figure 1: An example of a multitask network consisting of $N = 11$ nodes and $Q = 3$ clusters. Each color represents the nodes of a particular cluster. The solid lines are the connections between the nodes of the same cluster while the dashed lines represent the connections between the nodes of different clusters.

be collected into a $M \times 1$ column vector $\boldsymbol{\omega}^o \triangleq \text{col}\{\alpha_1, \alpha_2, \cdots \alpha_M\}$ and the past data into a $1 \times M$ regression vector:

$$\boldsymbol{x}_k^T(i) \triangleq [d_k(i-1), d_k(i-2), \cdots, d_k(i-M)]. \tag{3}$$

Now we can observe that the physical problem in (2) can be rewritten as the linear regression model in (1). Interested readers can find more examples in [36].

Contrary to the classic single task scenarios where all the agents have the same parameter vector to estimate (i.e. $\boldsymbol{\omega}_k^o = \boldsymbol{\omega}^o$ for all the nodes in the network), here we consider a multitask case where different clusters of agents have different objectives. We assume that there are Q different clusters in the network where $\mathcal{C}_q$ represents the nodes that belong to cluster q. The nodes of a cluster follow the same objective:

$$\boldsymbol{\omega}_k^o = \boldsymbol{\omega}_{\mathcal{C}_q}^o, \text{ for all } \quad k \in \mathcal{C}_q. \tag{4}$$

Note that the nodes of different clusters can be connected and share information with each other without having any prior clustering information. Figure 1 represents an example of such a network. In this network, there are $N = 11$ nodes in the network belonging to $Q = 3$ different clusters. In this figure, the solid lines are representative of connections between the nodes in the same cluster while the dashed lines represent the link between the

nodes of different clusters. In our model, we consider a general case where some of the nodes (such as node 3 in Figure 1) might be loosely connected to the nodes of their cluster. This means that a large portion of the information they receive can be useless and misleading for their own estimation task. Next, we formulate the distributed optimization task over the network to estimate the unknown parameters and introduce the multitask diffusion adaptation strategy to tackle this problem.

2.2. Multitask Diffusion Adaptation Strategy

To formulate the multitask diffusion strategy, we associate a local cost function, $J_k(\boldsymbol{\omega}_{\mathcal{C}(k)})$, with each node k where $\mathcal{C}(k)$ represents the cluster that node k belongs to:

$$J_k(\boldsymbol{\omega}_{\mathcal{C}(k)}) = \mathbb{E}\left\{|d_k(i) - \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_{\mathcal{C}(k)}|^2\right\}. \tag{5}$$

Now, the global cost function which is the aggregation of the nodes' cost functions can be formulated as:

$$J^{\text{glob}}(\boldsymbol{\omega}_{\mathcal{C}(1)}, ..., \boldsymbol{\omega}_{\mathcal{C}(Q)}) = \sum_{k=1}^{N} \mathbb{E}\left\{|d_k(i) - \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_{\mathcal{C}(k)}|^2\right\}. \tag{6}$$

It has been shown that the optimization problem in (6) is strongly convex and second-order differentiable [11]. Therefore, to solve this problem in a distributed manner, we employ the adapt-then-combine (ATC) multitask diffusion strategy [11]:

$$\boldsymbol{\psi}_k(i) = \boldsymbol{\omega}_k(i-1) + \mu_k[d_k(i) - \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_k(i-1)]\boldsymbol{x}_k(i), \tag{7}$$

$$\boldsymbol{\omega}_k(i) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k}(i)\boldsymbol{\psi}_\ell(i), \tag{8}$$

where $\boldsymbol{\psi}_k(i)$ is the intermediate estimate of node k at time i, $\mu_k \geq 0$ is the updating step-size and $\boldsymbol{\omega}_k(i)$ is the estimation of node k for $\boldsymbol{\omega}_k^o$ at time i. The weights $a_{\ell k}(i)$ in (8) are called *combination weights* and as can be seen in Figure 2, each $a_{\ell k}(i)$ is the weight that node k assigns to the information received from node ℓ at time instant i. The coefficients $a_{\ell k}(i)$ are the non-negative elements of the $N \times N$ combination matrix $\boldsymbol{A}_i$ for each time instant i. Moreover, the combination weights $a_{\ell k}(i)$ must satisfy:

$$\sum_{\ell \in \mathcal{N}_k} a_{\ell k}(i) = 1,\ a_{\ell k}(i) > 0 \text{ if } \ell \in \mathcal{N}_k,\ a_{\ell k}(i) = 0 \text{ if } \ell \notin \mathcal{N}_k. \tag{9}$$

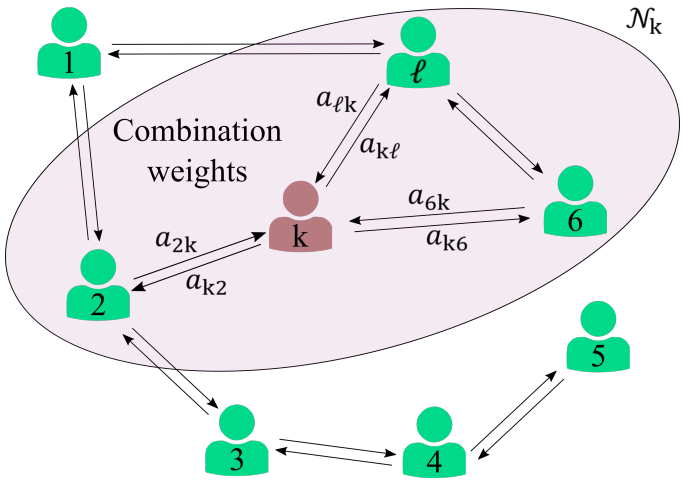

Figure 2: An example of a connected network where the neighboring nodes share information with each other. The combination weight $a_{\ell k}(i)$ is the weight that node k assigns to the information received from node ℓ at time instant i.

It can be observed from equation (8) that in the combination step of the diffusion strategy each node k combines the intermediate estimates of its neighbors $\boldsymbol{\psi}_\ell(i)$ to obtain its own estimate at time instant i. However, in the more realistic and challenging scenario where agents are reluctant to share their genuine information with all their neighbors, the received information from a neighboring node ℓ might be different from its true estimate $\boldsymbol{\psi}_\ell(i)$. In our model, genuine information of each node k refers to its true estimate $\boldsymbol{\psi}_k(i)$, which is obtained according to equation (7), and sharing anything rather than $\boldsymbol{\psi}_k(i)$ is considered as fabricated information. Now, we need to integrate the information credibility issue and the multitask diffusion concepts. We propose a more general formulation of the ATC strategy to incorporate this idea:

$$\boldsymbol{\psi}_k(i) = \boldsymbol{\omega}_k(i-1) + \mu_k[d_k(i) - \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_k(i-1)]\boldsymbol{x}_k(i), \tag{10}$$

$$\boldsymbol{\psi}_\ell^{\mathrm{rec},k}(i) = \begin{cases} \boldsymbol{\psi}_\ell(i), & k \in \mathcal{N}_{\mathbb{G}}^{\ell}(i) \\ \boldsymbol{\chi}_\ell(i), & \text{otherwise} \end{cases} \tag{11}$$

$$\boldsymbol{\omega}_k(i) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k}(i)\boldsymbol{\psi}_\ell^{\mathrm{rec},k}(i), \tag{12}$$

where $\boldsymbol{\psi}_{\ell}^{\mathrm{rec},k}(i)$ is the information received by node k from node ℓ at time instant i, and $\boldsymbol{\chi}_{\ell}(i)$ is the $M \times 1$ fabricated information shared by node ℓ at time instant i. This fabricated data can be anything rather than the true estimate $\boldsymbol{\psi}_k(i)$ and can be of various forms such as Gaussian or ~~Chi~~ square distributions. Moreover, $\mathcal{N}_{\mathbb{G}}^{\ell}(i)$ represents the subset of node ℓ's neighbors that node ℓ shares its genuine information with at time instant i. Since node k is not aware whether the received information is genuine or fabricated, it combines all the received information $\boldsymbol{\psi}_{\ell}^{\mathrm{rec},k}(i)$ from its neighbors via the combination weights according to equation (12).

It should be noted that there are several ways to design the combination weights. It has been shown in [12, 30] that the selection of combination weights $a_{\ell k}(i)$ has a significant impact on the performance of multitask networks. As mentioned earlier, the neighboring nodes might have different objectives which means that they could be exposed to information that is not related to their own objective. Therefore, the combination weights must be designed in a way that helps the nodes to ignore this information that might misdirect them from their task. Hence, it is important to design the combination weights such that they assign a greater weight to neighbors with similar objectives and lower weights to neighbors from different clusters. It has been shown that designing the weights to minimize the instantaneous mean-square deviation (MSD) of the network (which is a metric that measures the error variance of the agents) results in ~~the~~ optimal combination weights ~~as follows~~ [42]:

$$\mathrm{MSD}(i) \triangleq \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}\|\widetilde{\boldsymbol{\omega}}_k(i)\|^2, \tag{13}$$

where $\widetilde{\boldsymbol{\omega}}_k(i) \triangleq \boldsymbol{\omega}_k^o - \boldsymbol{\omega}_k(i)$ is the error vector at node k at iteration i and $\|\boldsymbol{x}\|$ represents the Euclidean norm of the vector argument $\boldsymbol{x}$. Then, the combination coefficients $a_{\ell k}(i)$ can be obtained by solving the optimization problem:

$$\min_{\{a_{\ell k}(i)\}} \mathrm{MSD}(i) = \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}\|\widetilde{\boldsymbol{\omega}}_k(i)\|^2. \tag{14}$$

The optimal solution can be approximated by [42, 36, 12]:

$$a_{\ell k}(i) \approx \begin{cases} \frac{\|\boldsymbol{\omega}_k(i-1) - \boldsymbol{\psi}_{\ell}^{\mathrm{rec},k}(i)\|^{-2}}{\sum_{n \in \mathcal{N}_k} \|\boldsymbol{\omega}_k(i-1) - \boldsymbol{\psi}_{n}^{\mathrm{rec},k}(i)\|^{-2}}, & \ell \in \mathcal{N}_k \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

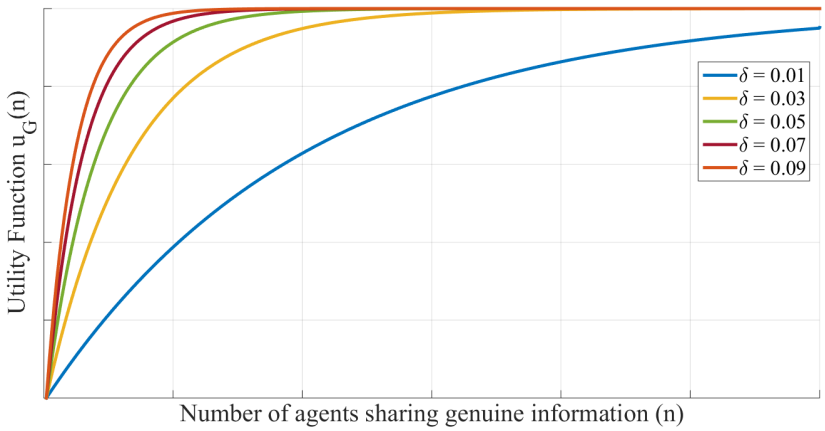

Figure 3: The utility function $u_{\mathbb{G}}(n)$ for different values of δ.

One important observation from (15) is that the combination weights are estimated such that they act as a similarity measure; the nodes allocate higher weights to neighbors with similar objectives while learning to ignore the deceptive information of others. Using this combination method enables the nodes to continuously learn about the objective of their neighbors so that they can distinguish the useful and irrelevant information. Estimating the combination weights in this manner helps the agents to benefit from cooperation with their neighbors in multitask scenarios. However, there are several cases where some of the nodes decide to share fabricated information instead of their genuine data even with those in the same cluster. In the next section, we model and study this behavior to be able to propose a method to overcome this problem.

3. Information Credibility Modeling

As mentioned earlier, agents of a cooperative network can share information with each other. However, sharing their true information can be costly in several cases. As a result, the agents might decide to manipulate the information they share due to privacy concerns or to obtain unfair benefits for themselves. Here, we formulate the information credibility problem in multitask diffusion networks and investigate its effect on the performance of distributed learning methods. In order to model this behavior, we follow the approach in [20] and assume that at the beginning of each time instant i, each agent can decide whether to share information truthfully or manipulate

and share fabricated misleading information. These two strategies can be represented by $\mathbb{G}$ and $\mathbb{F}$, respectively. Clearly, this decision by each node k depends on the cost that it has to pay for sharing genuine information, which can be represented by c_k. Moreover, the probability of sharing genuine information for agent k at time instant i, represented by $p^i_{k,\mathbb{G}}$, depends on this cost.

Since the nodes are connected to, and collaborating with each other the utility of each node depends on the actions and behaviors of other nodes in the network. In other words, we can assume that the benefit of utilizing shared information increases as the number of genuine agents in the network grows. Here we employ a general utility function that has been used widely for distributed networks [7, 20]:

$$u^k_{\mathbb{G}}(n) = \gamma - e^{-\delta n} - c_k, \tag{16}$$

where $u^k_{\mathbb{G}}(n)$ is the utility function for agent k sharing genuine information, and n is the number of agents sharing genuine information at a specific time instant. Moreover, $\gamma, \delta \geq 0$ are constant parameters where δ represents the speed of saturation for genuine information and γ is the maximum utility of the user. It is important to note that the values of the parameters $0 \leq \gamma \leq 1$ and c_k are normalized in our model and can be directly compared to each other in this manner. Although the cost of sharing genuine information c_k can vary among the nodes of the network, we assume that c_k is the same for all agents and can be replaced by c without loss of generality. Figure 3 shows the characteristics of the utility function for different values of δ. On the other hand, the utility function of an agent k sharing false information, $u^k_{\mathbb{F}}(n)$, can be formulated as:

$$u^k_{\mathbb{F}}(n) = p_a(\gamma - e^{-\delta(n+1)}) + (1 - p_a)(\gamma - e^{-\delta n}). \tag{17}$$

It should be noted that p_a is the probability of acquiring new information for each agent, which decreases as the information acquisition cost increases. Moreover, the term $n+1$ in the first exponent is due to the n genuine agents in the network plus node k's own genuine data that it has access to.

Now that the utility function for both genuine and false agents is formulated, we can obtain the average utility for sharing genuine information $\bar{u}_{\mathbb{G}}(p_{\mathbb{G}})$ over the network:

$$\bar{u}_{\mathbb{G}}(p_{\mathbb{G}}) = \sum_{n=1}^{N-1} \binom{N-1}{n} p^n_{\mathbb{G}}(1 - p_{\mathbb{G}})^{N-1-n}(\gamma - e^{-\delta(n+1)} - c). \tag{18}$$

In a similar way, the average utility for sharing false information, $\bar{u}_{\mathbb{F}}(p_{\mathbb{G}})$, can be obtained by:

$$\bar{u}_{\mathbb{F}}(p_{\mathbb{G}}) = \sum_{n=1}^{N-1} \binom{N-1}{n} p_{\mathbb{G}}^{n}(1-p_{\mathbb{G}})^{N-1-n}(p_a(\gamma - e^{-\delta(n+1)}) + (1-p_a)(\gamma - e^{-\delta n})). \tag{19}$$

Therefore, the average utility function of the whole network can be formulated as:

$$\bar{u}(p_{\mathbb{G}}) = p_{\mathbb{G}}\bar{u}_{\mathbb{G}}(p_{\mathbb{G}}) + (1-p_{\mathbb{G}})\bar{u}_{\mathbb{F}}(p_{\mathbb{G}}). \tag{20}$$

Having defined the utility functions across the network, the probability of sharing genuine information $p_{\mathbb{G}}$ at each time instant i can be obtained by a differential equation using evolutionary game theory [15, 20]:

$$p_{\mathbb{G}}^{i+1} = p_{\mathbb{G}}^{i} + \alpha[\bar{u}_{\mathbb{G}}(p_{\mathbb{G}}^{i}) - \bar{u}(p_{\mathbb{G}}^{i})], \tag{21}$$

where α is a constant positive variable. It can be observed from (21) that in the cases where the utility of sharing genuine information is higher than the average utility, the probability of sharing genuine information increases over the agents enticing them to be more truthful. To observe the dynamics of the information sharing process, we show the effect of the information sharing cost c on the probability of sharing genuine information $p_{\mathbb{G}}^{i}$ in Figure 4. It can be seen from this figure that increasing the information sharing cost c results in a decrease of probability of sharing genuine information over the agents of the network. This means that there is a trade-off between the cost for sharing genuine information and the chance of receiving truthful and beneficial data among the nodes. Since a low probability of sharing genuine information can result in a degraded benefit of cooperation and the poor performance of the network, it is crucially important to develop a strategy to overcome this problem. In the following section, we propose a reputation protocol to enhance the performance of the adaptive network.

4. Adaptive Distributed Reputation Protocol

As discussed earlier, the agents of an adaptive distributed network might be inclined to share false and misleading information rather than their genuine data due to several reasons. In Section 3 we studied this behavior and modeled the dynamics of probability of sharing genuine information among

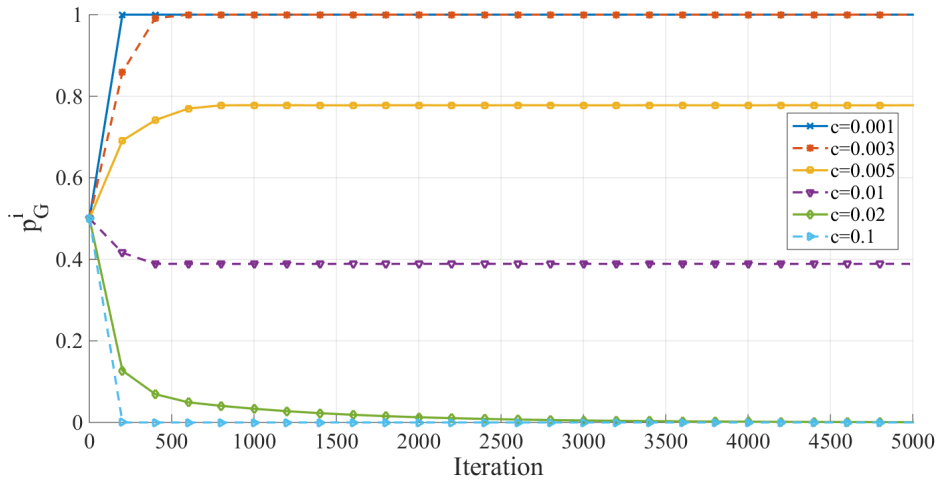

Figure 4: The probability of sharing genuine information $p_{\mathbb{G}}^i$ for different values of cost c.

the agents. It was observed that several factors such as the high cost of information sharing can contribute to a low probability of sharing truthful information. Therefore, it is important to tackle this obstacle by an adaptive and distributed strategy. Here we propose a reputation scoring method which enables the agents to evaluate the importance of their neighbors for their own estimation task. With the help of this reputation scheme, each agent can summarize the past actions of its neighbors as a reputation score. This score is then used to evaluate the importance and truthfulness of each particular neighbor. Using this score, the agents are then able to select a subset of their neighbors that are more beneficial and contributive to them for their own estimation task. As a result, each agent shares its genuine information with this subset of neighbors while sharing false information with the remaining neighbors according to the probability of sharing genuine information $p_{\mathbb{G}}^i$.

4.1. Reputation Score

Several reputation strategies have been proposed and used in the literature to entice cooperative behavior [18, 34, 9, 21]. According to reputation protocols, cooperative and truthful agents have higher reputation scores among the agents while non-cooperative and deceptive agents suffer form a lower reputation score [2]. Among several mechanisms to design reputation scores (such as average and cumulative scores), an exponentially-weighted

moving average scheme is more robust to cheating. This is due to the fact that the exponentially-weighted scores allocate higher weights to the recent observations compared to the past actions [18, 40]. For the first time, we propose to utilize such reputation scores in multitask diffusion strategies to help the agents to overcome the destructive effects of misleading information. To design the reputation scores, we follow a similar exponentially-weighted approach that has been proposed in [18]. However the reputation scores here are based on the dynamics of sharing genuine information and the importance and similarity of the received data.

Therefore, we formulate the cooperative reputation score as a smooth exponential moving average which allocates a higher weight to recent activities and a lower weight to the history of the nodes. The reputation score $R_{\ell k}(i)$ that agent k assigns to its neighbor ℓ at time i can therefore be formulated as:

$$R_{\ell k}(i) = \beta_k R_{\ell k}(i-1) + (1-\beta_k) a_{\ell k}(i) \tag{22}$$

where $\beta_k \in (0,1)$ is the smoothing factor that node k uses to control the evolution of the reputation scores. It can be observed from (22) that a higher value of β_k results in a higher weight and influence of the past actions.

Each node k of the network can then utilize the obtained reputation scores $R_{\ell k}(i)$ of its neighbors at each time instant to evaluate the importance of each particular neighbor for its own estimation task. Using this score, each node k can sort its neighbors so that the nodes with higher reputation scores have a higher rank. Based on this ranking and by considering the probability of sharing genuine information, node k forms a subset of its neighbors to share genuine information with. In particular, node k shares its own genuine information with a subset of its neighbors that have higher reputation scores. These are the nodes that are more beneficial and contributive for the estimation task of node k. The dynamics of this subset can change over time according to the dynamics of the probability of sharing genuine information $p_{\mathbb{G}}^i$. To make sure that none of the agents violate the probability $p_{\mathbb{G}}^i$, the maximum number of the agents that each node k can share genuine information with at each time instant should be less than or equal to:

$$\mathcal{N}_{\mathbb{G},k}^{\max}(i) = \lfloor p_{\mathbb{G}}^i \times |\mathcal{N}_k| \rfloor \tag{23}$$

where $\mathcal{N}_{\mathbb{G},k}^{\max}(i)$ represents the maximum number of neighbors that node k at time instant i can share genuine information with. Moreover, $|\mathcal{N}_k|$ is the cardinality of the neighbor set $\mathcal{N}_k$ and the floor function $\lfloor x \rfloor$ maps the

Algorithm 1 Summary of the proposed multitask genuine information sharing algorithm.

Require: $\boldsymbol{\omega}_k(0) = \boldsymbol{\psi}_k(0)$, for all k

Set the values to model information credibility: $p^0_{\mathbb{G}}, \gamma, \delta, p_a, c, \alpha$

for $i \geq 1$ **do**

Step 1: Probability of sharing genuine information-[Section 3]

$\bar{u}_{\mathbb{G}}(p_{\mathbb{G}}) = \sum_{n=1}^{N-1} \binom{N-1}{n} p_{\mathbb{G}}^n (1-p_{\mathbb{G}})^{N-1-n} (\gamma - e^{-\delta(n+1)} - c)$ [ref. (18)]

$\bar{u}_{\mathbb{F}}(p_{\mathbb{G}}) = \sum_{n=1}^{N-1} \binom{N-1}{n} p_{\mathbb{G}}^n (1-p_{\mathbb{G}})^{N-1-n} (p_a(\gamma - e^{-\delta(n+1)}) + (1-p_a)(\gamma - e^{-\delta n}))$ [ref. (19)]

$\bar{u}(p_{\mathbb{G}}) = p_{\mathbb{G}} \bar{u}_{\mathbb{G}}(p_{\mathbb{G}}) + (1-p_{\mathbb{G}}) \bar{u}_{\mathbb{F}}(p_{\mathbb{G}})$ [ref. (20)]

$p^{i+1}_{\mathbb{G}} = p^i_{\mathbb{G}} + \alpha[\bar{u}_{\mathbb{G}}(p^i_{\mathbb{G}}) - \bar{u}(p^i_{\mathbb{G}})]$ [ref. (21)]

Step 2: Multitask diffusion strategy-[Section 2.2]

$\boldsymbol{\psi}_k(i) = \boldsymbol{\omega}_k(i-1) + \mu_k[d_k(i) - \boldsymbol{x}_k^T(i)\boldsymbol{\omega}_k(i-1)]\boldsymbol{x}_k(i)$, [ref. (10)]

$$\boldsymbol{\psi}_\ell^{\mathrm{rec},k}(i) = \begin{cases} \boldsymbol{\psi}_\ell(i), & k \in \mathcal{N}^\ell_{\mathbb{G}}(i) \\ \boldsymbol{\chi}_\ell(i), & \text{otherwise} \end{cases}$$ [ref. (11)]

$$a_{\ell k}(i) \approx \begin{cases} \frac{\|\boldsymbol{\omega}_k(i-1) - \boldsymbol{\psi}_\ell^{\mathrm{rec},k}(i)\|^{-2}}{\sum_{n \in \mathcal{N}_k} \|\boldsymbol{\omega}_k(i-1) - \boldsymbol{\psi}_n^{\mathrm{rec},k}(i)\|^{-2}}, & \ell \in \mathcal{N}_k \\ 0, & \text{otherwise} \end{cases}$$ [ref. (15)]

$\boldsymbol{\omega}_k(i) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k}(i) \boldsymbol{\psi}_\ell^{\mathrm{rec},k}(i)$, [ref. (12)]

Step 3: Reputation Strategy-[Section 4]

$R_{\ell k}(i+1) = \beta_k R_{\ell k}(i) + (1-\beta_k) a_{\ell k}(i)$ [ref. (22)]

$\mathcal{N}^{\max}_{\mathbb{G},k}(i) = \lfloor p^i_{\mathbb{G}} \times |\mathcal{N}_k| \rfloor$ [ref. (23)]

$\mathcal{N}^k_{\mathbb{G}}(i+1) =$ [the first $\mathcal{N}^{\max}_{\mathbb{G},k}(i)$ neighbors of node k that have higher reputation scores $R_{\ell k}(i)$].

end for

real number x to the largest previous integer. ow, at each time instant i each agent k ranks its neighbors according to thNeir reputation scores $R_{\ell k}(i)$ obtained by (22) and share its genuine information with the first $\mathcal{N}_{\mathbb{G},k}^{\max}(i)$ neighbors that have higher reputation scores. For example, if the probability of sharing genuine information at time i is $p_{\mathbb{G}}^{i} = 0.5$, agent k shares its genuine information with at most half of its neighbors that have higher reputation scores. We will show in Section 5 that by selecting the subset of genuine agents in this manner, the multitask network has better performance compared to other learning strategies. Moreover, it is observed that the performance of the proposed method is comparable to the ideal and unrealistic case where all the agents share genuine information. The summary of the proposed algorithm is given in Algorithm 1.

5. Numerical Experiments

In this section, we evaluate the performance of the proposed method and compare it with those of several other diffusion strategies in the case where the agents are subject to a cost for sharing genuine information. As discussed earlier, this cost could be due to several reasons such as privacy preservation, security constraints, or obtaining an unfair bias. This cost results in a selfish behavior among the nodes and can entice them to share false or fabricated data rather than their true information. Utilizing the probability of sharing genuine information formulated in (21), we studied the performance of different strategies and compared them with the proposed adaptive reputation based method summarized in Algorithm 1.

5.1. Simulation Results

Here, the performance of five different learning strategies are evaluated:

(a) The non-cooperative method where each agent of the network tries to estimate its own parameter vector without receiving or sharing any information with the other agents. Therefore, the combination matrix $\boldsymbol{A}_i = \boldsymbol{I}_N$, which means that there is no aggregation of information from the neighbors in the combination step of the diffusion strategy.

(b) The standard diffusion adaptation algorithm with uniform combination weights $a_{\ell k}(i) = \frac{1}{|\mathcal{N}_k|}$. In this case, each node allocates the same weight to the information received from each of its neighbors without considering the cluster they belong to.

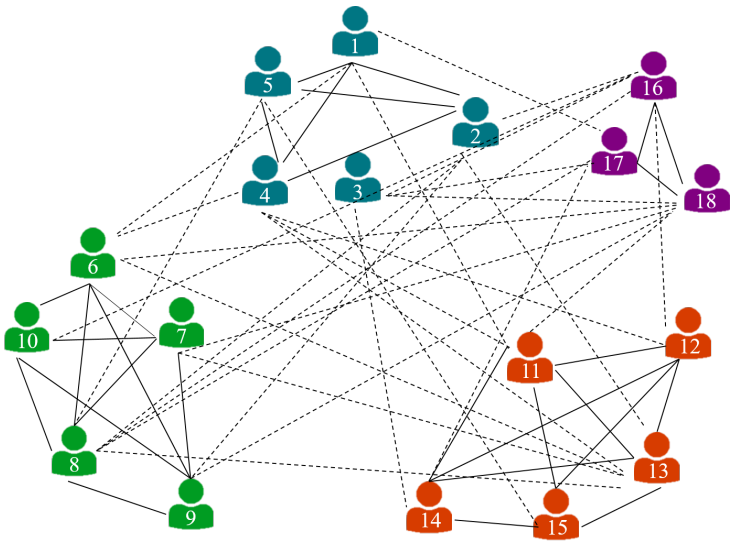

Figure 5: Topology of multitask network used in the simulation. The network consists of $N = 18$ nodes and $Q = 4$ clusters. The solid lines shows the connections between the nodes of the same cluster while the dashed lines represent the connections between the nodes of different clusters.

(c) The adaptive multitask method without the reputation scheme. In this case, the combination weights are estimated according to equation (15) but each node k selects the subset $\mathcal{N}_{\mathbb{G}}^{k}(i)$ for sharing genuine information in a random manner rather than using the reputation scores.

(d) The proposed adaptive multitask reputation-based method where both the combination weights and the subsets $\mathcal{N}_{\mathbb{G}}^{k}(i)$ are selected in the adaptive manner proposed in Algorithm 1.

(e) The adaptive multitask algorithm where all the agents share their true information with their neighbors. Obviously, this is an ideal and unrealistic scenario where all the agents share their genuine data despite its cost. However, we present the results of this unrealistic case only as a benchmark for comparison purposes. Moreover, we compare the aggregated communication cost of the network in this case with the other cases that take into account the cost and probability of sharing genuine information.

In the first simulation setup, we consider a network of $N = 18$ nodes which consists of $Q = 4$ different clusters: $C_1 = \{1, 2, 3, 4, 5\}$, $C_2 = \{6, 7, 8, 9, 10\}$, $C_3 = \{11, 12, 13, 14, 15\}$, and $C_4 = \{16, 17, 18\}$. The agents are connected over an exogenously determined topology shown in Figure 5. The regression

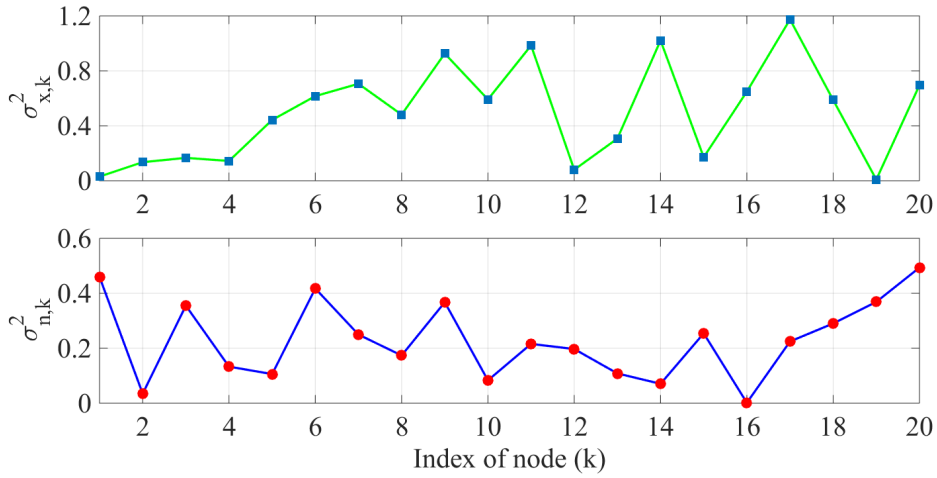

Figure 6: Variances of the input regression vectors (top) and noise (bottom) for each node of the network.

Network				**Info. Credibility**			**Reputation**			
N	Q	M	μ	c	$p^0_{\mathbb{G}}$	p_a	γ	δ	α	β
18	4	2	0.02	0.01	0.5	0.8	1	0.1	1	0.9

Table 2: Simulation parameters used for the simulation setup in Section 5.1.

input signals $\boldsymbol{x}_k(i)$ are zero-mean Gaussian random vectors of size $M \times 1$ where $M = 2$. The covariance matrices of the regression signals are considered to be diagonal and are obtained by $\boldsymbol{R}_{x,k} = \sigma^2_{x,k}\boldsymbol{I}_M$. Moreover, the measurement noise $n_k(i)$ has a Gaussian distribution with zero-mean and variance $\sigma^2_{n,k}$. The values of $\sigma^2_{x,k}$ and $\sigma^2_{n,k}$ for different nodes of the network are shown in Figure 6. The unknown parameter vector for each cluster is of size $M \times 1$ and is chosen as: $\boldsymbol{\omega}^o_{\mathcal{C}_1} = [0.5, -0.4]^T$, $\boldsymbol{\omega}^o_{\mathcal{C}_2} = [-1, 3]^T$, $\boldsymbol{\omega}^o_{\mathcal{C}_3} = [5.2, 2.8]^T$, and $\boldsymbol{\omega}^o_{\mathcal{C}_4} = [-2.2, 4.8]^T$. Moreover, the updating step-size μ_k of the diffusion algorithm in (10) is assumed to be the same for all the nodes of the network and is set to $\mu_k = \mu = 0.02$. The values of the parameters for the adaptive reputation protocol are $\gamma = 1$, $\delta = 0.1$, $p_a = 0.8$, $\alpha = 1$, and $\beta_k = \beta = 0.9$. Here, the communication cost c is set to 0.01. Table 2 shows the values of the parameters for the simulation setup in Section 5.1. To study the behavior of different strategies, we evaluate their performance for a wide range of conditions such as different network typologies and different distributions of the misleading information $\boldsymbol{\chi}_k(i)$.

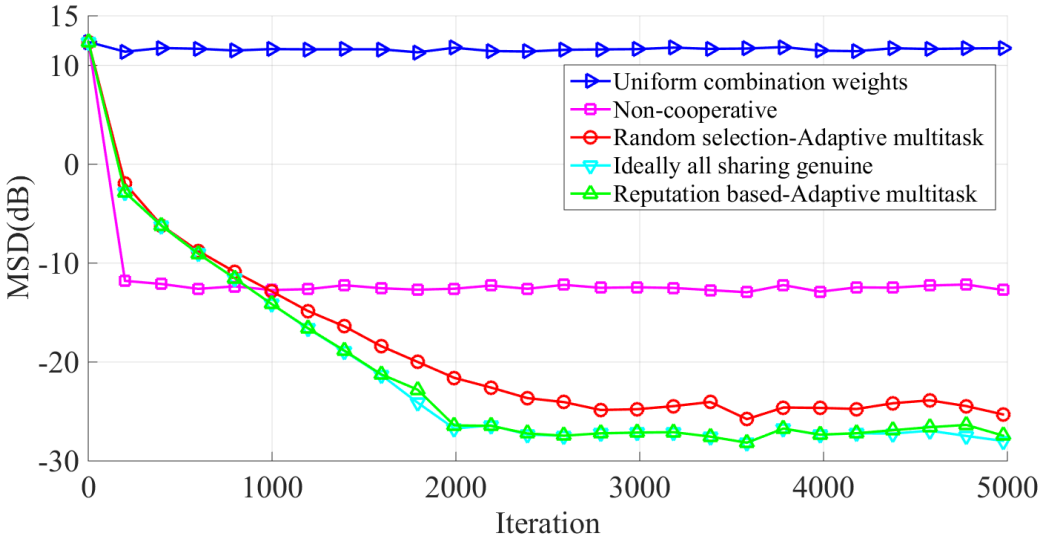

Figure 7: The MSD curves representing the performance of different learning strategies. Here the cost c is equal to 0.01 and the misleading data shared by the nodes follow a Gaussian distribution. Lower values of MSD show lower estimation error and better performance.

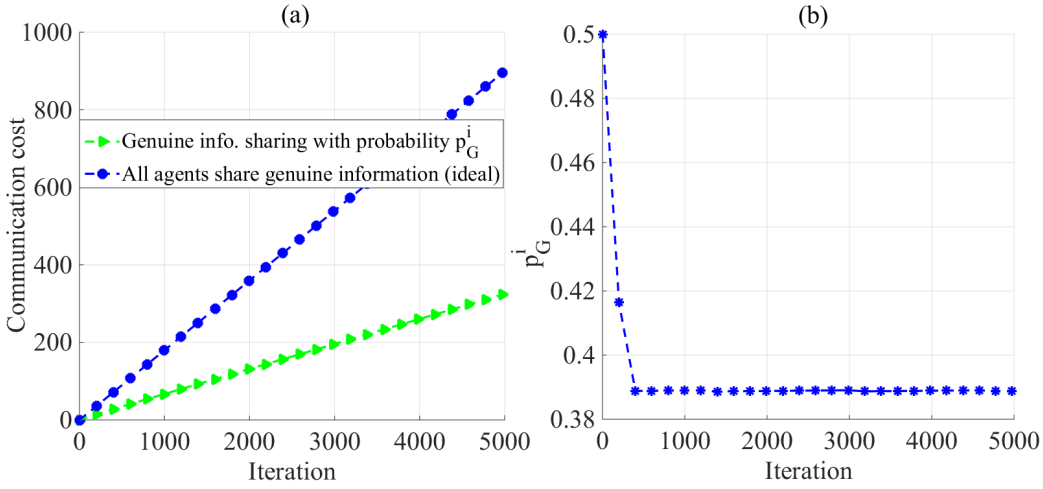

Figure 8: (a) The aggregated communication cost obtained by multiplying the number of data units shared genuinely over the network by the cost of sharing a genuine information unit c at each time instant, and (b) the probability of sharing genuine information over the network where the cost of sharing a genuine information unit c is equal to 0.01.

In the first setting, we assume that the fabricated information $\boldsymbol{\chi}_k(i)$ shared by node k is drawn randomly from a Gaussian distribution with probability density $\mathcal{N}((1,1)^T, \boldsymbol{I}_M)$. Figure 7 shows the MSD curves of different strategies for this case and the curves are obtained by averaging over 100 Monte-Carlo runs. As introduced in equation (13), MSD curves represent the error of the nodes in estimating their objectives, therefore, lower values of MSD mean lower estimation error and better performance. The MSD curves in this figure show that the proposed adaptive reputation-based method outperforms all the other learning strategies as it has a lower MSD. Here, it is interesting to note that the performance of the diffusion method with uniform combination weights is even worse than the case where there is no cooperation. This is due to the fact that by allocating uniform weights to all the neighbors, the nodes are not able to distinguish the nodes from different clusters; therefore, most of them are not successful in their estimation tasks. Moreover, we can observe that the performance of the proposed method is quite similar to that of the ideal case where all the agents share their true information over time. This implies that with the help of the proposed reputation strategy, the nodes are able to adapt and overcome the adverse effects of the fabricated data. Especially, this reveals the advantage of the proposed method when comparing its aggregated communication cost with the ideal scenario shown in Figure (8)(a). This figure represents the aggregated communication cost over the network.To obtain this cost for the case where the agents share their true information with probability $p_{\mathbb{G}}^i$, we calculate the number of data units shared genuinely over the network and multiply it by the cost of sharing a genuine information unit c at each time instant. ~~Then,~~ this cost is aggregated over time to obtain the ~~whole~~ communication cost spent by the network. Obviously, in the unrealistic case where all the shared information is genuine, the number of all the shared information units is multiplied by cost c at each time instant and the communication cost is much higher as the nodes fail to comply with the probability $p_{\mathbb{G}}^i$. In this case, the nodes violate the privacy and security constraints of the network while in the proposed method all the nodes behave in accordance with the existing costs and restrictions. Moreover, Figure 8(b) shows the dynamics of the probability of sharing genuine information $p_{\mathbb{G}}^i$. As can be observed from this figure, the initial probability is set to 0.5 and it decreases over time according to equation (21).

Figure (9) shows the evolution of a sample node (node 1) selecting a subset of neighbors with whom to share genuine information at 500 time in-

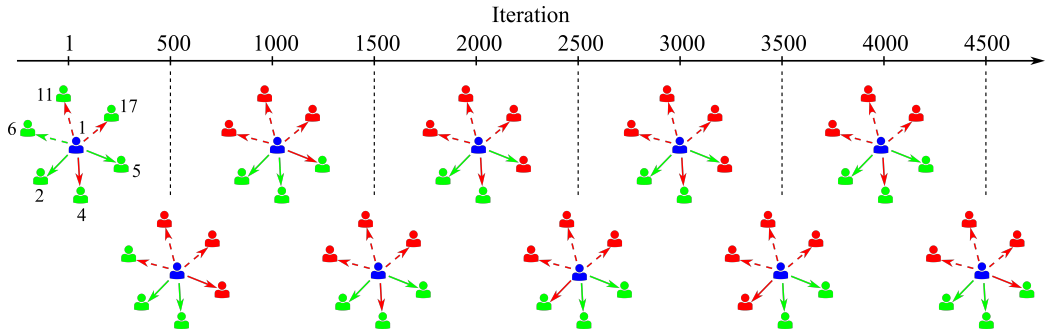

Figure 9: Evolution of node 1 selecting a subset of neighbors with whom to share genuine information.The green arrows represent the nodes receiving genuine information from node 1, while red ones show fabricated data from node 1. The color of the neighboring nodes at each iteration shows whether they are sending genuine information to node 1 or not.

tervals. In this figure, the green arrows represent the nodes that ~~are receiving~~ genuine information from node 1, while red ones show fabricated data from node 1. Additionally, the color of the neighboring nodes at each iteration shows whether they are sending genuine information to node 1 or not. For instance, we observe that at iteration 1000 nodes $\{2, 4, 5\}$ are sending genuine information to node 1, while nodes $\{6, 11, 17\}$ are sending fabricated data. It should be noted that this information is not available to node 1 and this node selects the subset $\mathcal{N}_{\mathbb{G}}^{1}(i)$ using the reputation scores according to (22).

In the previous simulation setup, it is assumed that the fabricated data follows a Gaussian distribution. However, in the real-world scenarios the fabricated or misleading data $\boldsymbol{\chi}_k(i)$ might not be drawn from any specific distribution such as the Gaussian distribution. Therefore, it is important to evaluate the performance of the different methods for more general cases where the nodes deal with other types of fabricated data. In order to do so, we generate the misleading information $\boldsymbol{\chi}_k(i)$ using chi-square distributions with various degrees of freedom ν. Figure 10 represents the result where $\boldsymbol{\chi}_k(i)$ is drawn from a chi-square distribution with degree of freedom $\nu = 3$. The other parameters such as the sharing cost c remain the same. Therefore, the dynamics of $p_{\mathbb{G}}^{i}$ and the aggregated communication cost of the network is the same as that in Figure 8.

The results in Figure 10 reveal that the proposed reputation-based method is robust to various forms of fabricated data while the performance of other leaning strategies might be affected by it. In particular, it can be observed

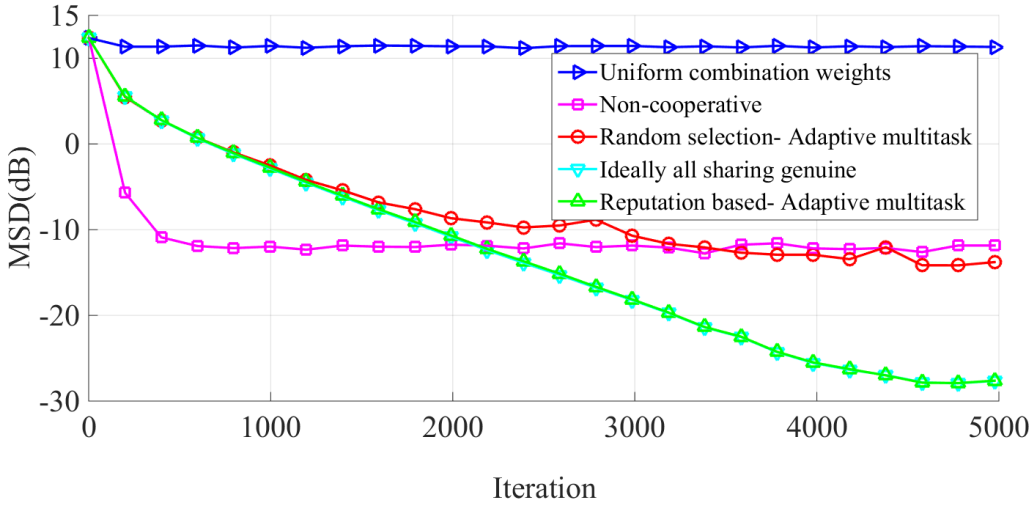

Figure 10: The MSD curves representing the performance of different learning strategies over the distributed network. Here the cost c is equal to 0.01 and the misleading data shared by the nodes follow a chi-square distribution. Lower values of MSD show lower estimation error and better performance.

from this figure that the performance of the adaptive multitask algorithm where the subset $\mathcal{N}_{\mathbb{G}}^{k}(i)$ is randomly selected is adversely affected and is almost the same as the non-cooperative case.

5.2. Large Arbitrary Networks

In order to study the performance of the proposed method in more probabilistic cases, we obtained the results for several arbitrary networks with large number of agents. Here we present the results of one of these cases, where a connected network consists of $N = 100$ nodes in $Q = 12$ randomly selected clusters where the size of clusters varies from 5 to 18 nodes. Moreover, two arbitrary nodes are connected to each other with a probability of 0.07. The objectives $\boldsymbol{\omega}_{\mathcal{C}_q}^{o}$ of the clusters are of size 2×1 and its entries are shown in Figure 11. The other parameters of this simulation setup is the same as those in Table 2. Moreover, the fabricated information $\boldsymbol{\chi}_k(i)$ shared by node k is drawn randomly from a Gaussian distribution with probability density $\mathcal{N}((1,1)^T, \boldsymbol{I}_M)$. Figure 12 shows the performance of the proposed method as well as the other learning strategies over the large network with

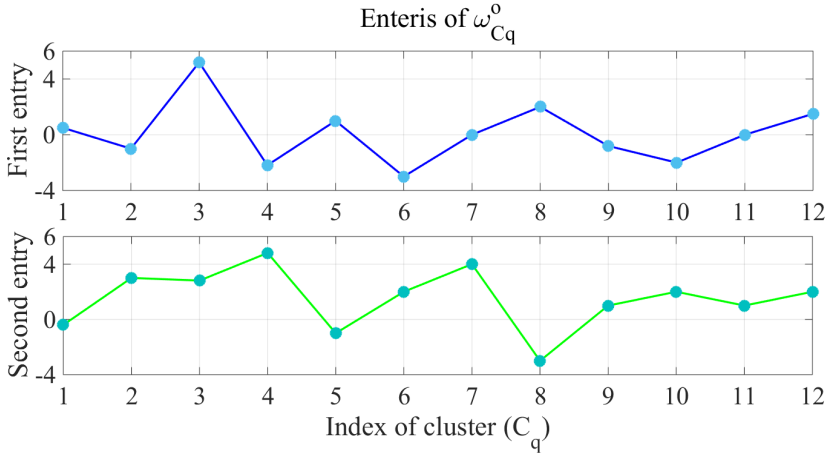

Figure 11: The objectives $\boldsymbol{\omega}^{o}_{\mathcal{C}_q}$ of the 12 different clusters. The objectives are of size 2×1, where the upper figure shows its first entry and the lower figure shows its second entry.

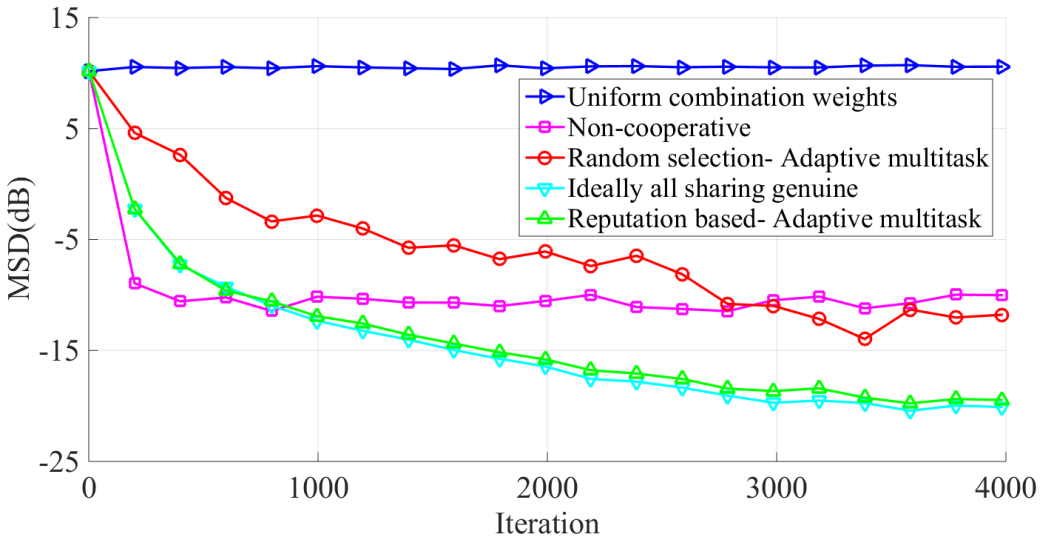

Figure 12: The MSD curves representing the performance of different learning strategies over a large distributed network with $N = 100$ nodes and 12 randomly selected clusters. Here the cost c is equal to 0.01 and the misleading data shared by the nodes follow a Gaussian distribution. Lower values of MSD show lower estimation error and better performance.

arbitrary topology. It can be observed from this figure that our proposed reputation based method has a lower MSD which means that it outperforms the other learning strategies. Moreover, its performance is quite close to the unrealistic case where all the agents are genuine ~~showing~~ the power of the adaptive reputation strategy.

6. Conclusion

In this paper, we studied the information credibility over multitask distributed networks where sharing true information is restricted due to privacy and security constraints. Therefore, the agents of the network are only able to share their true data with a subset of their neighbors rather than all of them. We proposed an adaptive reputation strategy which enables the agents to choose this subset of neighbors based on the role they play in their own estimation task. The results showed that our method has a superior performance compared to the other learning methods. Moreover, it was observed that the proposed method is robust to various forms of misleading and fabricated data. This is especially important for the scenarios where the cost of sharing information is high and the nodes are reluctant to share their true data to preserve their security.

References

[1] D. Acemoglu and A. Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49, 2011.

[2] D. Artz and Y. Gil. A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):58–71, 2007.

[3] S. Ba and P. A. Pavlou. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS quarterly*, pages 243–268, 2002.

[4] D. Bergemann and D. Ozmen. Optimal pricing with recommender systems. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 43–51. ACM, 2006.

[5] D. P. Bertsekas. A new class of incremental gradient methods for least squares problems. *SIAM Journal on Optimization*, 7(4):913–926, 1997.

[6] D. Blatt, A. O. Hero, and H. Gauchman. A convergent incremental gradient method with a constant step size. *SIAM Journal on Optimization*, 18(1):29–51, 2007.

[7] M. Bouakiz and M. J. Sobel. Inventory control with an exponential utility criterion. *Operations Research*, 40(3):603–608, 1992.

[8] P. Braca, S. Marano, and V. Matta. Enforcing consensus while monitoring the environment in wireless sensor networks. *IEEE Transactions on Signal Processing*, 56(7):3375–3380, 2008.

[9] S. Braynov and T. Sandholm. Contracting with uncertain level of trust. *Computational Intelligence*, 18(4):501–514, 2002.

[10] L. Canzian, Y. Xiao, W. Zame, M. Zorzi, and M. Van der Schaar. Intervention with private information, imperfect monitoring and costly communication. *IEEE Transactions on Communications*, 61(8):3192–3205, 2013.

[11] J. Chen, C. Richard, and A. H. Sayed. Multitask diffusion adaptation over networks. *IEEE Transactions on Signal Processing*, 62(16):4129–4144, 2014.

[12] J. Chen, C. Richard, and A. H. Sayed. Diffusion LMS over multitask networks. IEEE Trans. Signal Process., 63(11):2733–2748, 2015.

[13] J. Chen and A. H. Sayed. Diffusion adaptation strategies for distributed optimization and learning over networks. *IEEE Transactions on Signal Processing*, 60(8):4289–4305, 2012.

[14] J. Chen and A. H. Sayed. Distributed pareto-optimal solutions via diffusion adaptation. In *Proceedings of the IEEE Statistical Signal Processing Workshop (SSP)*, pages 648–651, 2012.

[15] R. Cressman. *Evolutionary dynamics and extensive form games*, volume 5. MIT Press, 2003.

[16] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[17] C. Dellarocas. Reputation mechanism design in online trading environments with pure moral hazard. *Information Systems Research*, 16(2):209–230, 2005.

[18] M. Fan, Y. Tan, and A. B. Whinston. Evaluation and design of online cooperative feedback mechanisms for reputation management. *IEEE Transactions on Knowledge and Data Engineering*, 17(2):244–254, 2005.

[19] O. N. Gharehshiran, V. Krishnamurthy, and G. Yin. Distributed energy-aware diffusion least mean squares: Game-theoretic learning. *IEEE Journal of Selected Topics in Signal Processing*, 7(5):821–836, 2013.

[20] C. Jiang, Z. Han, Y. Ren, and L. Hanzo. Information credibility equilibrium of cooperative networks. In *Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2016.

[21] A. Jøsang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision support systems*, 43(2):618–644, 2007.

[22] R. M. Karp. *Reducibility among combinatorial problems*. Springer, 1972.

[23] F. Li, B. Luo, P. Liu, D. Lee, and C.-H. Chu. Enforcing secure and privacy-preserving information brokering in distributed information sharing. *IEEE transactions on information forensics and security*, 8(6):888–900, 2013.

[24] C. G. Lopes and A. H. Sayed. Incremental adaptive strategies over distributed networks. *IEEE Transactions on Signal Processing*, 55(8):4064–4077, 2007.

[25] C. G. Lopes and A. H. Sayed. Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Transactions on Signal Processing*, 56(7):3122–3136, 2008.

[26] J. K. MacKie-Mason and H. R. Varian. Pricing congestible network resources. *IEEE journal on Selected Areas in Communications*, 13(7):1141–1149, 1995.

[27] S. Monajemi, K. Eftaxias, S. Sanei, and S.-H. Ong. An informed multi-task diffusion adaptation approach to study tremor in parkinson's disease. *IEEE Journal of Selected Topics in Signal Processing*, 10(7):1306–1314, 2016.

[28] S. Monajemi, S. Ensafi, S. Lu, A. A. Kassim, C. L. Tan, S. Sanei, and S.-H. Ong. Classification of HEp-2 cells using distributed dictionary learning. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, 2016.

[29] S. Monajemi, S. Sanei, and S.-H. Ong. Advances in bacteria motility modelling via diffusion adaptation. In *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, 2014.

[30] S. Monajemi, S. Sanei, S.-H. Ong, and A. H. Sayed. Adaptive regularized diffusion adaptation over multitask networks. In *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–5. IEEE, 2015.

[31] A. Nedic and D. P. Bertsekas. Incremental subgradient methods for non-differentiable optimization. *SIAM Journal on Optimization*, 12(1):109–138, 2001.

[32] A. Nedic and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.

[33] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[34] P. Resnick and R. Zeckhauser. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. In *The Economics of the Internet and E-commerce*, pages 127–157. Emerald Group Publishing Limited, 2002.

[35] A. H. Sayed. Diffusion adaptation over networks. *Academic Press Library in Signal Processing*, 3:323–454, 2013.

[36] A. H. Sayed. Adaptation, learning, and optimization over networks. *Foundations and Trends in Machine Learning*, 7(4-5):311–801, 2014.

[37] J. N. Tsitsiklis and M. Athans. Convergence and asymptotic agreement in distributed decision problems. *IEEE Transactions on Automatic Control*, 29(1):42–50, 1984.

[38] S-Y. Tu and A. H. Sayed. Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks. *IEEE Transactions on Signal Processing*, 60(12):6217–6234, 2012.

[39] J. Xu, Y. Song, and M. Van Der Schaar. Sharing in networks of strategic agents. *IEEE Journal of Selected Topics in Signal Processing*, 8(4):717–731, 2014.

[40] C.-K. Yu, M. van der Schaar, and A. H. Sayed. Information-sharing over adaptive networks with self-interested agents. *IEEE Transactions on Signal and Information Processing over Networks*, 1(1):2–19, 2015.

[41] Y. Zhang, J. Park, and M. van der Schaar. Rating protocols in online communities. *ACM Transactions on Economics and Computation*, 2(1):4, 2014.

[42] X. Zhao and A. H. Sayed. Clustering via diffusion adaptation over networks. In *Proc. 3rd International Workshop on Cognitive Information Processing (CIP)*, pages 1–6, 2012.

[43] D. Zissis and D. Lekkas. Addressing cloud computing security issues. *Future Generation computer systems*, 28(3):583–592, 2012.