On the Privacy Guarantees of Gossip Protocols in General Networks

Richeng Jin, *Member, IEEE*, Yufan Huang, Zhaoyang Zhang, *Senior Member, IEEE*, Huaiyu Dai, *Fellow, IEEE*

***Abstract*—Recently, the privacy guarantees of information dissemination protocols have attracted increasing research interests, among which the gossip protocols assume vital importance in various information exchange applications. In this work, we study the privacy guarantees of gossip protocols in general networks in terms of differential privacy and prediction uncertainty. First, lower bounds of the differential privacy guarantees are derived for gossip protocols in general networks in both synchronous and asynchronous settings. The prediction uncertainty of the source node given a uniform prior is also determined. For the *private gossip* algorithm, the differential privacy and prediction uncertainty guarantees are derived in closed forms in the asynchronous setting. Moreover, considering that these two metrics may be restrictive in some scenarios, the relaxed variants are proposed. It is found that source anonymity is closely related to some key network structure parameters in the general network setting. Then, we investigate information spreading in wireless networks with unreliable communications, and quantify the tradeoff between differential privacy guarantees and information spreading efficiency. Finally, considering that the attacker may not be present at the beginning of the information dissemination process, the scenario of delayed monitoring is studied and the corresponding differential privacy guarantees are evaluated.**

***Index Terms*—Information spreading, gossip protocols, differential privacy, prediction uncertainty.**

I. Introduction

It is well-known that most people are six or fewer social connections away from each other. Recently, the explosive development of the Internet and social networks makes it easy for people to disseminate their information to the rest of the world. Gossip protocols, in which networked nodes randomly choose a neighbor to exchange information, have been widely adopted in various applications for information dissemination due to their simplicity and efficiency. For instance, they can be used to spread and aggregate information in dynamic networks like mobile networks, wireless sensor networks, and unstructured P2P networks [2]–[5]. Combined with stochastic gradient descent methods, gossip protocols are also adapted to implement distributed machine learning [6], [7]. In particular, the authors of [7] propose to transmit differentially private gradient information through gossip protocols. Nonetheless, they focus on the privacy of the shared gradient information rather than the anonymity of the source.

With the arising concerns of privacy exposure, information sources often prefer to stay anonymous while disseminating sensitive information. Gossip protocols are believed to provide a certain form of source anonymity since most nodes do not get informed directly by the source, and the origin of the information becomes increasingly blurred as the spreading proceeds. In this regard, source identification and protection of gossip protocols have attracted significant research interests (e.g., see [8], [9] and the references therein). Various criteria and estimators have been developed to determine the source node, including rumor centrality [10]–[12], Jordan centrality [13]–[15], Gaussian source estimator [16]–[18], Markov Chain Monte Carlo-based estimator [19], Bayesian source estimator [20] and maximum a posteriori estimator [21]. Instead of focusing on source identification, [9], [22], [23] proposed adaptive diffusion protocols to hide the source node. However, the existing approaches usually assume some specific network structures (e.g., tree graphs) and attacking techniques (e.g., maximum likelihood estimator), and therefore don't easily generalize.

To study the privacy of gossip protocols in a formal and rigorous setting, the concept of differential privacy [24], which was originally introduced in data science, is adapted to measure the source anonymity of gossip protocols in [25]. However, their study is restricted to complete networks, which may not be a good model in practice. For example, practical networks often have a network diameter much larger than 1 (41 for the Facebook network [26]). In [1], we study the fundamental limits on the privacy of gossip-based information-spreading protocols in general networks. In this work, we further derive the privacy guarantees of the private gossip algorithm [25] in closed form and present extensive numerical results on various network structures. In addition, we propose two new metrics, i.e., the candidate-set-based differential privacy and prediction uncertainty, to better capture the practical privacy concerns of the source node. More specifically, our main contributions are summarized as follows.

1) Lower bounds of the differential privacy guarantees of general gossip protocols are derived for general networks in both synchronous and asynchronous settings. The prediction uncertainty of the source node given a uniform prior is also determined.
2) For the private gossip algorithm [25], the differential privacy guarantees and prediction uncertainty are derived in closed form. In addition, considering that the

R. Jin and Z. Zhang are with the Zhejiang–Singapore Innovation and AI Joint Research Lab, the Department of Information and Communication Engineering, Zhejiang University, Hangzhou, China, 310007, and also with Zhejiang Provincial Key Lab of Information Processing, Communication and Networking (IPCAN), Hangzhou 310007, China (e-mail: {richengjin, ning_ming}@zju.edu.cn). Y. Huang and H. Dai are with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA, 27695 (e-mail: {yhuang20, hdai}@ncsu.edu). Part of this work was presented at 2020 IEEE Global communications conference [1].

original differential privacy and prediction uncertainty may be restrictive in some scenarios, candidate set-based differential privacy and prediction uncertainty variants are proposed. Numerical results are presented to show the privacy guarantees of the private gossip algorithm in different network structures.

3) The privacy guarantees of standard gossip and private gossip protocols are further studied in a wireless setting, where communications are assumed to be unreliable. It is found that wireless interference can enhance the differential privacy while slowing down the spreading process. Through analysis and simulations, the tradeoff between the differential privacy guarantees and the information spreading efficiency is revealed.
4) The effect of the additional uncertainty induced by delayed monitoring on the privacy guarantees is shown.

The remainder of this work is organized as follows. Section II discusses the related work. Section III introduces the system model. The privacy guarantees of gossip protocols are presented in Section IV. The privacy-spreading tradeoff of gossip protocols in wireless networks is discussed in Section V. Section VI investigates the privacy guarantee of gossip protocols in the delayed monitoring scenario. Conclusion and future works are discussed in Section VII.

II. Related Work

Gossip protocols are widely adopted to model the spreading of infectious diseases [27] and information dissemination in dynamic networks [2]–[5]. On the one hand, in terms of epidemic spreading [28], it is essential to accurately identify the source nodes. In this sense, locating the source of a gossip mechanism has attracted significant research efforts (e.g., see [8] and the references therein). Various criteria have been proposed to determine the source node. [10], [11] proposed rumor centrality which measures the number of distinct propagation paths originating from the source, and the node with the maximum rumor centrality is identified as the source node in regular trees. [12] designated a set of nodes as suspicious sources and considered the node with the maximum rumor centrality in the set of the suspicious sources as the source node. [13]–[15] proposed Jordan centrality which considers the source node associated with the optimal sample path (which is the path that most likely leads to the observations of the entity that intends to identify the source node) as the propagation origin. Various estimators have also been proposed, such as, Gaussian source estimator [16]–[18], Markov chain Monte Carlo-based estimator [19], Bayesian source estimator [20], maximum a posteriori estimator [21] and maximum likelihood estimator [29], [30]. [31] proposed a types-center method that serves as an approximation of the maximum-likelihood source estimator. [32] introduced a Gaussian weighted averaging correlation coefficient to evaluate the likelihood of possible sources. [33] proposed a latent space mapping-based method for source identification, and [34] proposed a machine learning-based source identifier for social networks. Different from most of the aforementioned techniques that consider static networks, [35]–[39] studied the source identification problem in time-varying networks, and [40] studied the source identification in multiplex networks. [41] considered both truth and rumor diffusions and identified the sources of truth and rumor simultaneously. [42] considered source identification in the process of contact tracing in epidemic spreading.

On the other hand, in terms of information spreading on social media platforms, it may be preferred to protect the anonymity of the message authors. To the best of our knowledge, there are only a few works on preserving the anonymity of the source node. [9], [22], [23] proposed adaptive diffusion protocols to hide the source node. However, the privacy of the source node is only guaranteed in tree graphs. Moreover, it is recently shown that the adaptive diffusion protocols fail to protect the source node when the adversary has access to multiple observations [43]. [44] showed that the source node can escape detection by modifying the network structure. In this sense, it is of particular interest to quantify the privacy of gossip protocols with a formal privacy notion and investigate the impact of network structures.

Recently, differential privacy [24] has emerged as a strong candidate for privacy measure due to its strong information-theoretic guarantees. It finds applications in a wide variety of areas, such as location privacy preservation [45], privacy-aware data release [46], and privacy-preserving deep learning [47], to name a few. [25] first adopted the concept of differential privacy to study the privacy guarantees of gossip protocols and proposed the private gossip algorithm. Nonetheless, it only considered the complete network graph. In this work, we study the privacy guarantees of gossip protocols in general networks and investigate the impact of imperfect communications among the nodes and delayed observation from the adversary.

III. System Model

In this section, we first introduce the considered gossip protocols and the time model in Section III-A and Section III-B, respectively. Then, we introduce the threat model that describes the capability of the adversary in Section III-C. Finally, the formal privacy notions are detailed in Section III-D. Important notations used in this article are summarized in Table I.

A. Gossip Protocol

Given a connected network $G = (V, E)$ of arbitrary topology, where $V = \{0, 1, ..., n-1\}$ is the node set and E is the set of connecting edges, a node (source) initially possesses a piece of information and needs to deliver it to all the other nodes in the network. All the nodes are assumed to share the same communication protocol $gossip$. Each time an informed node i performs $gossip$, it will contact one of its neighboring nodes $j \in N_i$ uniformly at random (i.e., with probability $1/d_i$, where d_i is the degree of node i). The whole information dissemination process terminates after all the nodes are informed. Same as [25], we focus on the gossip protocols based on the "push"

TABLE I: Important notations

G	the connected network.
V	the set of nodes.
n	the number of nodes.
d_i	the degree of node i.
I	the set of the informed nodes.
A^c	the set of active nodes
α	the probability that the activities of the active nodes are observed by the attacker.
$\mathcal{S}$	the event observed by the attacker.
$\mathcal{D}^{(i)}$	the source indicator that the i-th node is the source.
$p_G^{(i)}(\mathcal{S})$	the conditional probability of the observation event $\mathcal{S}$ given the network graph G and the source node i.
(ϵ, δ)	the differential privacy parameters.
c	the prediction uncertainty.
$p_G(I_0\|\mathcal{S})$	the posterior probability of the source node being I_0 given the observation event $\mathcal{S}$.
$\mathcal{Q}$	the candidate set of the source node.
$C_\beta(i)$	the decay centrality of node i.
D_G	the diameter of the graph.
$A[j,i]$	the probability of the active node j contacting node i.
A^m	the m-th power of A.
$P(j \to i)$	the probability of secret message spreading from source node j to another node i.
f	the failure probability of the communications between two nodes.
C_G	the cover time of a random walk in network G.
T_{as}	the spreading time of standard gossip given perfect communication.

action[1] in this work, and consider the following two specific gossip protocols, which are given in Algorithm 1.

1) Standard Gossip: All informed nodes remain active (i.e., continuously performing *gossip*) during the spreading process.
2) Private Gossip [25]: Once an active informed node (initially it is the source) performs *gossip*, it turns inactive, and the newly informed node takes over the source role.

Algorithm 1 Gossip Algorithms

```
Require: The number of nodes $n$, the source node $k$.
Ensure: The informed node set $I = \{0, 1, \cdots, n-1\}$
Initialization: Informed node set $I \leftarrow \{k\}$, active node set $A^c \leftarrow \{k\}$
while $|I| < n$ do
    The active node in $A^c$ performs gossip and another node $j$ is informed
    $\begin{cases} I \leftarrow I \cup \{j\}, A^c \leftarrow A^c \cup \{j\}, & \text{for standard gossip,} \\ I \leftarrow I \cup \{j\}, A^c \leftarrow \{j\}, & \text{for private gossip.} \end{cases}$
end while
```

B. Time Model

Both synchronous and asynchronous time models are considered. In the former, all nodes share a global discrete time clock. Each time the clock ticks, all active informed nodes perform the *gossip* action simultaneously, and the informed node set is updated accordingly, counted as one round. In the asynchronous time model, each node has its own internal clock, which ticks according to a Poisson process, with the mean interval between two ticks equivalent to that of one round in the synchronous model. The *gossip* action and update of the informed node set is performed each time the clock of an active informed node ticks.

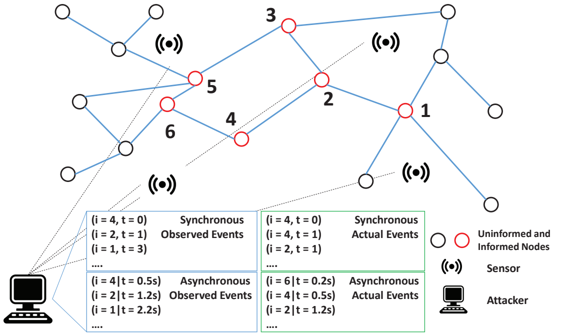


Fig. 1: Sensor Monitoring and Observations

C. Threat Model

The goal of the attacker is to identify the source node based on its observations (i.e., attack on confidentiality and privacy). It is assumed that the attacker can monitor the ongoing communications in the whole network, through, e.g., deploying a sufficient number of sensors throughout the field. With a probability of $0 < \alpha \leq 1$, the sensors can correctly observe the identities of the active nodes at each gossip step. Specifically, as shown in Fig. 1, the observed event has the form of $\mathcal{S} = ((i,t))$, $i \in V$, $t \in \{0, 1, 2, \cdots\}$ in the synchronous setting, which indicates that the attacker knows node i performs the gossip action at time slot t. In the asynchronous setting, however, the attacker does not know the exact time of each observed event, but only the relative order of the nodes' activities. The observed event in this case is represented by $\mathcal{S} = ((i|t))$, where the condition t stands for the latent time information unknown to the attacker.

D. Privacy Model

In this work, differential privacy is adopted to measure the information leakage of the gossip protocols. In particular, a randomized algorithm $\mathcal{R}$ with domain $\mathbb{N}^{|\mathcal{X}|}$ is (ϵ, δ)-differentially private if for all $\mathcal{S} \subseteq Range(\mathcal{R})$ and for any two databases x, y that differ on a single element [24]:

$$Pr[\mathcal{R}(x) \in \mathcal{S}] \leq e^{\epsilon} Pr[\mathcal{R}(y) \in \mathcal{S}] + \delta, \quad (1)$$

where parameter $\epsilon \geq 0$ is the privacy budget while $\delta \geq 0$ is the tolerance level for the violation of the ϵ bound. Specifically, given the privacy budget ϵ and the tolerance level δ, Eq. (1) implies that the randomized algorithm guarantees that the privacy loss is bounded by ϵ with a probability of at least $1 - \delta$. Intuitively, the smaller the ϵ and the δ, the better the privacy. Consider a source indicator database of the format $\mathcal{D}^{(i)} = [0, ..., d_i = 1, ..., 0]$ with exactly one nonzero value $d_i = 1$ if node i is the source. Given $D \triangleq \{\mathcal{D}^{(i)}\}_{i=0}^{n-1}$ and the

[1] In the corresponding "pull" action, uninformed nodes are active and try to solicit the information from informed nodes. The "push" action is dominant for information spreading in social and mobile networks. In addition, such a study is also conservative in the sense that it gives the attacker an advantage by only monitoring the "push" actions.

graph G as the input, a gossip protocol can be treated as a randomized algorithm with the output set $\mathbb{S}$ (i.e., the range) consisting of all possible observation sequences by the attacker during the execution of the protocol.

Definition 1. *[25] Given a general network G, a gossip protocol is (ϵ, δ)-differentially private in G if for all observations $\mathcal{S} \subseteq \mathbb{S}$ and for any two source indicator vectors $\mathcal{D}^{(i)}, \mathcal{D}^{(j)}, i, j \in V$:*

$$p_G^{(i)}(\mathcal{S}) \leq e^{\epsilon} p_G^{(j)}(\mathcal{S}) + \delta, \tag{2}$$

where $p_G^{(i)}(\mathcal{S}) = Pr[\mathcal{S}|G, \mathcal{D}^{(i)}]$ is the conditional probability of an observation event $\mathcal{S}$ given the network graph G and the source indicator vector $\mathcal{D}^{(i)}$.

In this work, considering the fact that, due to the topological and observation model constraints, there may exist some (rare) events $\mathcal{S}$ such that $p_G^{(j)}(\mathcal{S}) = 0$ (e.g., if $\mathcal{S}_{i,0}$ is the observed event that node i performs *gossip* at time 0 in the synchronous setting, then $p_G^{(j)}(\mathcal{S}_{i,0}) = 0, \forall j \neq i$), additional tolerance level is needed to ensure the privacy guarantees. Thus, for the privacy guarantees of general gossip protocols, we mainly focus on the study of the tolerance level δ. For the private gossip algorithm (i.e., Algorithm 1 with private gossip) in the asynchronous setting, we are able to derive the corresponding differential privacy level ϵ and give some analysis.

In addition to differential privacy, it is also desirable to study privacy guarantees of information dissemination protocols from a more pertinent perspective, i.e., source identification through prediction or detection. Reusing the above example, there always exist some events $\mathcal{S}$ such that $p_G^{(i)}(\mathcal{S}) > 0$ for some i but $p_G^{(j)}(\mathcal{S}) = 0, \forall j \neq i \in V$, which satisfy an arbitrary privacy budget ϵ with a tolerance level of δ (if $p_G^{(i)}(\mathcal{S}) \leq \delta$). However, the identity of the source (i.e., node i) can still be easily inferred. Therefore, it is further required that some *prediction uncertainty* be guaranteed for a given differentially private protocol, which is defined as [25]:

Definition 2. *[25] Given a general network G, the prediction uncertainty of a gossip protocol is defined for a uniform prior $p_G(I_0)$ on source nodes and any $i \in \{0, 1, ..., n-1\}$ as:*

$$\begin{aligned} c &= \min_{i, \mathcal{S} \subseteq \mathbb{S}} \left(\frac{p_G(I_0 \neq \{i\}|\mathcal{S})}{p_G(I_0 = \{i\}|\mathcal{S})} \right) \\ &= \min_{i, \mathcal{S} \subseteq \mathbb{S}} \left(\frac{1}{p_G(I_0 = \{i\}|\mathcal{S})} \right) - 1, \ \forall p_G^{(i)}(\mathcal{S}) > 0, \end{aligned} \tag{3}$$

where I_0 stands for the initial informed node set and its element represents the source node.

Remark 1. *The connection between prediction uncertainty and differential privacy is illustrated below. For any observation $\mathcal{S} \subseteq \mathbb{S}$, differential privacy guarantees that the posterior probabilities of the attacker observing $\mathcal{S}$ given different source nodes are bounded by (2), i.e., the source nodes do not change the distribution of the observations much. Prediction uncertainty guarantees that the posterior probability of node i being the source node given $\mathcal{S}$ satisfies $p_G(I_0 = \{i\}|\mathcal{S}) \leq p_G(I_0 \neq \{i\}|\mathcal{S})/c$, i.e., the observations do not change the attacker's knowledge about the source nodes much. Especially, because of the uniform prior $p_G(I_0)$, $\frac{p_G(I_0 \neq \{i\}|\mathcal{S})}{p_G(I_0 = \{i\}|\mathcal{S})} = \frac{\sum_{j \neq i} p_G^{(j)}(\mathcal{S})}{p_G^{(i)}(\mathcal{S})}$ holds by the Bayes' formula. Prediction uncertainty is an appealing metric in this study as it measures the privacy guarantees from the source prediction perspective with a much smaller cardinality than the classic privacy budget (which requires the study of all pairs of $p_G^{(i)}(\mathcal{S})$ and $p_G^{(j)}(\mathcal{S})$). Moreover, given a prediction uncertainty c, it can be shown that $p_G(I_0 = \{i\}|\mathcal{S}) \leq \frac{1}{c+1}, \forall i, \mathcal{S}$; therefore a larger c indicates better source anonymity.*

In order to further remedy the aforementioned issue of differential privacy, a relaxed differential privacy variant, termed differential privacy within a candidate set, is proposed. Its definition is given as follows.

Definition 3. *Given a general network G, a gossip protocol is (ϵ, δ)-differentially private within a candidate set $\mathcal{Q}$ in G if for all observations $\mathcal{S} \subseteq \mathbb{S}$ and for any two source indicator vectors $\mathcal{D}^{(i)}, \mathcal{D}^{(j)}, i, j \in \mathcal{Q} \subseteq V$:*

$$p_G^{(i)}(\mathcal{S}) \leq e^{\epsilon} p_G^{(j)}(\mathcal{S}) + \delta, \tag{4}$$

where $p_G^{(i)}(\mathcal{S}) = Pr[\mathcal{S}|G, \mathcal{D}^{(i)}]$ is the conditional probability of an observation event $\mathcal{S}$ given the network graph G and the source indicator vector $D^{(i)}$.

Remark 2. *The only difference between differential privacy and differential privacy within a candidate set is that the privacy guarantee is ensured for the source nodes falling in a candidate set $\mathcal{Q}$ instead of the whole network V. Note that the notion of differential privacy is highly conservative (considering the worst-case scenario). If there exists a node i such that $p_G^{(i)}(\mathcal{S}) = 0$, the probability of node i being the source is 0 from the attacker's perspective. In such a case, the attacker can always differentiate node i from other nodes, which means that the differential privacy guarantee will be violated with a high probability. However, on the one hand, it does not necessarily mean that the gossip protocols cannot provide effective source privacy protection. Particularly, it may be difficult for the attacker to distinguish any pairs of other nodes except node i in the network (e.g., $p_G^j(\mathcal{S}) = p_G^z(\mathcal{S}), \forall j \neq z \in V \setminus \{i\}$). On the other hand, in practice, it may not be necessary for the source node to hide itself in the whole network. Instead, it may be enough to make itself indistinguishable from a subset of the network (e.g., its neighbors). Our definition of differential privacy within a candidate set measures the privacy guarantee of the gossip protocols in such scenarios. Especially, as long as $p_G^{(j)}(\mathcal{S}) > 0, \forall j \in \mathcal{Q}$, $\delta = 0$ is always feasible for differential privacy within the candidate set $\mathcal{Q}$. In addition, when $\mathcal{Q} = V$, it is equivalent to the original definition of differential privacy. Therefore, it can be understood as a relaxed version of differential privacy.*

Similarly, the prediction uncertainty within a candidate set is defined as follows.

Definition 4. *Given a general network G, the prediction uncertainty of a gossip protocol within the candidate set $\mathcal{Q}$ is defined for a uniform prior $p_G(I_0)$ on source nodes over the* *candidate set Q as:*

$$c = \min_{i \in Q, \mathcal{S} \subseteq \mathbb{S}} \left(\frac{\sum_{j \neq i \in Q} p_G(I_0 = \{j\} | \mathcal{S})}{p_G(I_0 = \{i\} | \mathcal{S})} \right), \ \forall p_G^{(i)}(\mathcal{S}) > 0. \quad (5)$$

IV. Privacy of Gossip Protocols in General Networks

In this section, the privacy guarantees of gossip protocols are investigated. The intuition behind the privacy preservation capability of the gossip protocols is as follows. Given the threat model in Section III-C, there is a chance that active nodes perform *gossip* successfully and the attacker fails to catch it. If the first observed event of the attacker is that node i performs *gossip* at time $t = k$, it can conclude that the source node is within the k-hop neighborhood of node i. However, without additional information, the attacker cannot tell exactly which node among the k-hop neighbors of node i is the source. In the meantime, it becomes even more difficult for the attacker to identify the source node if it does not have access to the exact time t of the observed event (which corresponds to the asynchronous case). The identity of the source node is therefore better protected. In the following, we make the above intuition rigorous and present the formal privacy guarantees of the gossip protocols. More specifically, we first analyze the privacy guarantees of general gossip protocols in Section IV-A, followed by our analyses on the private gossip algorithm and the corresponding numerical illustration in Section IV-B and Section IV-C, respectively. Considering that the original differential privacy and prediction uncertainty may not be good measures for source anonymity, we further investigate differential privacy and prediction uncertainty within the candidate set in Section IV-D. For the general gossip protocols, general results concerning the lower bounds of the tolerance level δ and the upper bounds of the prediction uncertainty c are obtained. For the private gossip algorithm in the asynchronous setting, the pure version of differential privacy ($\delta = 0$) is feasible, and the corresponding differential privacy level ϵ and prediction uncertainty c are derived.

A. General Gossip Protocols

In this subsection, the privacy guarantees of general gossip protocols in general networks are studied. To facilitate our following analysis, we need the following lemma and the definition of *decay centrality*.

Lemma 1. *Given any gossip protocol in a graph G, let $\mathcal{S} \subseteq \mathbb{S}$ and there are two constants $w_G^{(i)}(\mathcal{S}), w_G^{(j)}(\mathcal{S})$ such that $p_G^{(i)}(\mathcal{S}) \geq w_G^{(i)}(\mathcal{S})$ and $p_G^{(j)}(\mathcal{S}) \leq w_G^{(j)}(\mathcal{S})$. If the gossip protocol satisfies (ϵ, δ)-differential privacy, then $\delta \geq \max_{\mathcal{S},i,j}(w_G^{(i)}(\mathcal{S}) - e^{\epsilon} w_G^{(j)}(\mathcal{S}))$.*

Lemma 1 readily follows from the definition of differential privacy; its proof is omitted in the interest of space.

Definition 5. *[48] Given a network G and a decay parameter β, $0 < \beta < 1$, the **decay centrality** of node i is defined as*

$$C_{\beta}(i) = \sum_{j \neq i} \beta^{d(i,j)}, \quad (6)$$

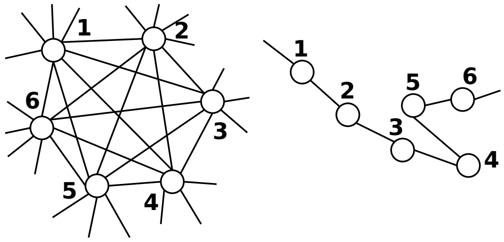


Fig. 2: Node 1 and node 6 are more distinguishable in the right network.

where $d(i,j)$ is the length of the shortest path between node i and j.

Remark 3. *Decay centrality measures the ease of a node reaching out to other nodes in the network. A large decay centrality indicates the central positioning of a node and its easiness to reach other nodes. The difficulty increases as β decreases.*

Our main result concerning the privacy guarantees of general gossip protocols in a general network is given below.

Theorem 1. *Given a connected network G with n nodes and diameter $D_G = \max_{i,j \in V, i \neq j} d(i,j)$, and considering the observation model described in Section III-C with parameter α, if a gossip protocol satisfies (ϵ, δ)-differential privacy for any $\epsilon \geq 0$ and c-prediction uncertainty, then we have $\delta \geq \alpha$ and $c = 0$ in the synchronous setting. In the asynchronous setting,*

$$\delta \geq \max \left[\alpha - e^{\epsilon}(1-\alpha)^{D_G}, \alpha - e^{\epsilon} \frac{1-\alpha}{n-1} \right] \quad (7)$$

and

$$c \leq \min_{i \in V} \frac{C_{1-\alpha}(i)}{\alpha}, \quad (8)$$

where $C_{1-\alpha}(i)$ is the decay centrality of node i with decay parameter $1 - \alpha$.

Proof: Please see Appendix A. ■

Remark 4. *Some interpretations of the results of Theorem 1 are in order. It can be observed that the asynchronous setting provides better privacy guarantees than the synchronous setting since the attacker has less information (i.e., the timing of the events) in this case. Note that differential privacy considers the worst-case scenario. In the synchronous setting, when the attacker detects the activity of a node at time 0, it can infer that the corresponding node is the source immediately. Therefore, the prediction uncertainty is 0 due to this worst-case event, and the privacy guarantees are determined by the attacker's sensing capability α in the synchronous setting. In the asynchronous setting, however, the attacker could not directly infer the source solely based on the first-observed event due to the lack of associated timing. A counter-example can be found in Fig. 1.*

As a result, the structure of the network plays an important role in the asynchronous setting. In the context of information spreading, if two nodes are further apart, it takes more time for the information to be spread from one to the other; this duration gives the attacker more opportunities to differentiate

This article has been accepted for publication in IEEE Transactions on Network Science and Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2023.3247626

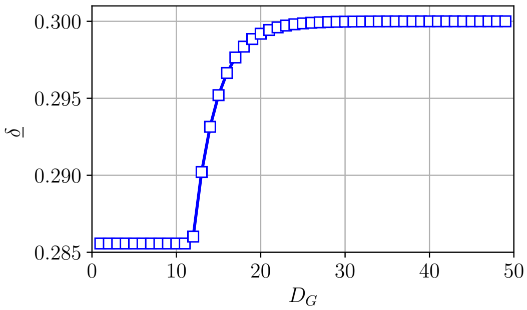


Fig. 3: Tolerance level v.s. network diameter ($\alpha = 0.3, n = 50, \epsilon = 0.01$).

the detected events, which leads to potentially higher privacy loss of the source node's identity. For instance, in the left network of Fig. 2, considering the event $\mathcal{S}_{1,0}$, i.e., node 1's activity is the first observed event by the attacker in the asynchronous setting, the probability of this event given that the source is 6 is $p_G^{(6)}(\mathcal{S}_{1,0}) \leq (1-\alpha)$ and the probability of this event given that the source is 1 is $p_G^{(1)}(\mathcal{S}_{1,0}) \geq \alpha$. But in the right network, the corresponding probabilities are $p_G^{(6)}(\mathcal{S}_{1,0}) \leq (1-\alpha)^5$ and $p_G^{(1)}(\mathcal{S}_{1,0}) \geq \alpha$, which makes $\mathcal{S}_{1,0}$ a more distinguishable event in the right network. Therefore, the network diameter D_G, as the distance measure of the whole network, captures the potential privacy loss and becomes a key factor of the differential privacy lower bound in (7); an example of the relationship between the differential privacy tolerance level bound and the network diameter is shown in Fig. 3. The same logic is reflected on the prediction uncertainty given in (8). The smaller the decay centrality a network has (i.e., the nodes are more distant from each other), the more likely the attacker can identify the source node through its observations. Therefore, the inherent network structure imposes a certain limit on privacy preservation concerning the source node identity, which applies to all information spreading protocols and calls for other privacy protection mechanisms, to be further explored in future work.

In addition, it can be seen that as the attacker's sensing capability α increases the privacy guarantees decrease (i.e., δ increases and c decreases). In particular, for an omnipresent attacker with $\alpha = 1$, we have $\delta = 1$ and $c = 0$ even in the asynchronous setting.

Finally, given (7), the lower bound of the differential privacy level ϵ in the asynchronous setting for a given δ can be obtained as follows

$$e^{\epsilon} \geq \max\left[\left(\frac{\alpha-\delta}{(1-\alpha)^{D_G}}\right), \left(\frac{(\alpha-\delta)(n-1)}{1-\alpha}\right)\right]. \tag{9}$$

B. Private Gossip Algorithm in the Asynchronous Setting

While the privacy guarantees of general gossip protocols are studied in Section IV-A, in this subsection, we investigate the private gossip algorithm (i.e., Algorithm 1 with private gossip) in more detail. Particularly, the differential privacy level ϵ and the prediction uncertainty c of the private gossip algorithm for general graphs in the asynchronous setting are examined, and closed-form expressions can be obtained owing to special characteristics of private gossip.[2]. In this case, the asynchronous setting is of more interest since there is no need for the nodes to coordinate and perform *gossip* simultaneously.

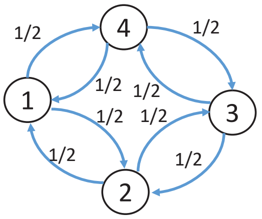


Fig. 4: The index of the active node inthe private gossip algorithm for a ring graph with 4 nodes. Each node has two neighbors and the probability of it contacting each neighboring node is $\frac{1}{2}$. For instance, if node 1 is active (i.e., the Markov chain is in state 1) at time $m-1$, then the probabilities of node 2 and node 4 being active at time m are both $\frac{1}{2}$.

In this case, since only one node is active at each time slot, the index of the active node can be modeled as a Markov chain with a transition probability matrix A. The (j,i)-th entry of A, denoted by $A[j,i]$, measures the probability of node j contacting node i once it becomes active. Fig. 4 shows an exemplary Markov chain of a ring graph with 4 nodes. In addition, to facilitate the discussion, we further define another matrix $\hat{A}_i$ as follows.

$$\hat{A}_i[j,k] = \begin{cases} A[j,k], & \text{if } j \neq i, \\ 0, & \text{if } j = i \text{ and } j \neq k, \\ 1, & \text{if } j = k = i. \end{cases} \tag{10}$$

Remark 5. *Note that $\hat{A}_i$ corresponds to the transition probability matrix of the Markov chain by setting node i as an absorbing state (i.e., node i will not push its message to any other nodes). Let $\hat{A}_i^m$ denote the m-th power of the matrix $\hat{A}_i$. Then, given the source node j, $\hat{A}_i^m[j,i]$ and $\hat{A}_i^{m-1}[j,i]$ are the probabilities of node i being active at time m and $m-1$, respectively. Note that if node i is active at time $m-1$, it is active at time m with probability 1 since it is an absorbing state. In this sense, given the source node j, the probability of node i being active for the first time at time m is $\hat{A}_i^m[j,i] - \hat{A}_i^{m-1}[j,i]$.*

To facilitate the discussion, we introduce the following quantity.

Definition 6. *The probability of **secret message spreading** from the source node j to another node i, denoted by $P(j \to i)$, is defined as the probability that, given the source node j, node i becomes active for the first time before the attacker observes the first event.*

Remark 6. *In the asynchronous setting, $P(j \to i)$ measures the similarity between node j and node i from the attacker's perspective. Intuitively, given source node j, if node i becomes active before the attacker observes the first event, the attacker cannot differentiate node j and node i based on its observation. The larger the $P(j \to i)$, the more difficult it is for the attacker to differentiate node j and node i.*

[2] We note that only one node is active at each time instance in the private gossip algorithm

This article has been accepted for publication in IEEE Transactions on Network Science and Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2023.3247626

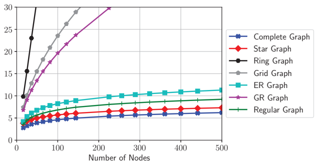


Fig. 5: ϵ v.s. number of nodes ($\alpha = 0.5$).

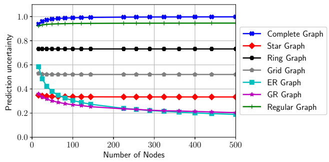


Fig. 6: c v.s. number of nodes ($\alpha = 0.5$).

With such consideration, the following lemma can be proved.

Lemma 2. *For the private gossip algorithm, given a general network G with source node j and the observation model described in Section III-C with parameter α, the probability of secret message spreading from the source node j to another node i for the first time is given by*

$$P(j \to i) = \alpha \sum_{m=0}^{\infty} (1-\alpha)^m \hat{A}_i^m[j,i] = \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i], \tag{11}$$

where I is an identity matrix and $\hat{A}_i$ is the transition probability matrix defined in (10).

Proof: Please see Appendix B. ■

Given Lemma 2, the privacy guarantees of the private gossip algorithm can be shown as follows.

Theorem 2. *Given a general network G and the observation model described in Section III-C with parameter α, the private gossip algorithm is $(\epsilon, 0)$-differentially private in the asynchronous setting, where*

$$\epsilon = \ln\left(\max_{j \neq i \in V} \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}\right). \tag{12}$$

The prediction uncertainty of the private gossip algorithm in the asynchronous setting is given by

$$c = \min_{i \in V} \sum_{j \neq i \in V} \alpha(I-(1-\alpha)\hat{A}_i))^{-1}[j,i]. \tag{13}$$

Proof: Please see Appendix C. ■

Remark 7. *Note that $\epsilon = \ln\left(\max_{j \neq i \in V} \frac{1}{P(j \to i)}\right)$ and $P(j \to i) = \alpha \sum_{m=0}^{\infty}(1-\alpha)^m \hat{A}_i^m[j,i]$. For any $m < d(j,i)$, it can be verified that $\hat{A}_i^m[j,i] = 0$. On the other hand, $(1-\alpha)^m$ decreases exponentially. In this sense, $P(j \to i)$ is supposed to decrease (and therefore ϵ will increase) as $d(j,i)$ increases, which verifies our discussion about the importance of the network structure (i.e., diameter) on differential privacy.*

In particular, given the $(\epsilon, 0)$-differential privacy guarantee, the corresponding tolerance level $\delta(\epsilon')$ can be obtained for any $\epsilon' > 0$ using the following corollary.

Corollary 1. *Any $(\epsilon, 0)$-differentially private mechanism is also $(\epsilon', \delta(\epsilon'))$-differentially private for all $\epsilon' > 0$, where*

$$\delta(\epsilon') = \Phi\left(-\frac{\epsilon'}{\mu_1} + \frac{\mu_1}{2}\right) - e^{\epsilon'}\Phi\left(-\frac{\epsilon'}{\mu_1} - \frac{\mu_1}{2}\right), \tag{14}$$

and $\mu_1 = -2\Phi^{-1}(\frac{1}{1+e^{\epsilon}})$.

Proof: Please see Appendix D. ■

C. Numerical Illustration

In this subsection, we examine the privacy guarantees of the private gossip algorithm numerically for several well-known graph topologies.[3]

Fig. 5 and Fig. 6 show the differential privacy level ϵ and prediction uncertainty c of the private gossip algorithm, respectively. In particular, in a complete graph, every node is connected to all the other nodes in the graph. A star graph is a graph in which a central node is connected to all the other nodes and the central node is their only neighbor. A ring graph is a graph that consists of a single cycle in which every node has exactly two edges incident with it. For the grid graph, we consider a two-dimensional square grid. These four graphs are deterministic (i.e., given the number of nodes, the graph structures are fixed). We also consider three random graphs, i.e., the regular graph, the Erdős Rényi (ER) graph, and the Geometric Random (GR) graph. An ER graph is a graph in which each node is randomly connected to another node with a certain probability. A GR graph is constructed by randomly placing the nodes in some metric space with the euclidean distance and connecting two nodes by an edge if and only if their distance is less than a specified parameter. A regular graph is a graph in which every node is randomly connected to a fixed number of nodes. In our simulation, we generate these three random graphs using the NetworkX package in Python such that the average degrees are 10.

It can be observed that as the number of nodes increases, the differential privacy level ϵ increases for all the graphs. On the one hand, in general, as the number of nodes increases, the probability of each node i being active given the source node j (and therefore $P(j \to i)$) decreases.[4]. On the other hand, the inequality in (2) should be satisfied for any pair of nodes. That being said, as the number of nodes increases, more inequalities need to be satisfied, which enforces a larger ϵ.

For prediction uncertainty, different graphs exhibit different trends. One possible reason is that for the complete graph and the regular graph, almost all the nodes are the same from the attacker's viewpoint. As a result, as the number of nodes increases, it becomes more difficult to identify the source node. On the other hand, since prediction uncertainty considers the worst-case scenario, as the number of nodes increases, the probability that there exists a node different from the other nodes (e.g., with a small degree) increases in the ER graph and the GR graph. As a result, the prediction uncertainty decreases as the number of nodes increases.

Given the same number of nodes, it can be seen from Fig. 5 and Fig. 6 that the regular graph and the complete graph (a special regular graph with the degree of $n-1$) perform well in terms of both differential privacy and prediction uncertainty. This may be due to the fact that in these graphs every node has the same degree while none of the nodes are distant from

[3] We note that unless otherwise noted, the differential privacy level and prediction uncertainty in the figures are the numerical evaluations of ϵ and c derived in the theorems.

[4] Note that in the private gossip, only one node is active at each time slot and the active node randomly selects its neighbors to push the message

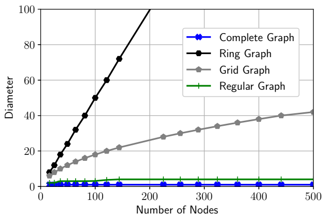


Fig. 7: Diameter v.s. the number of nodes.

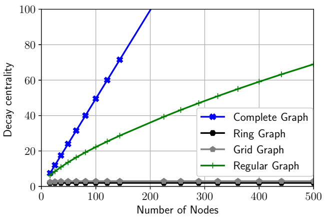


Fig. 8: Decay centrality v.s. number of nodes ($\alpha = 0.5$)

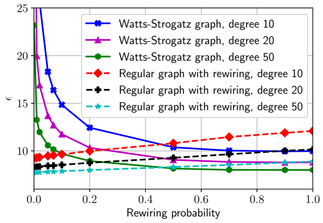


Fig. 9: ϵ v.s. rewiring probability (n = 500, $\alpha = 0.5$)

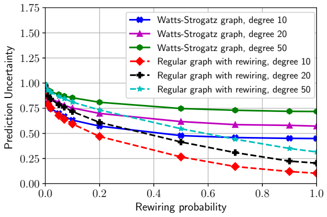


Fig. 10: c v.s. rewiring probability (n = 500, $\alpha = 0.5$)

the others. As a result, all the nodes look the same from the attacker's perspective. It is worth mentioning that despite that the ring graph and the grid graph can be considered as regular graphs with degrees of 2 and 4, respectively,[5] the diameters of these two graphs are large (which also leads to small decay centrality since a large diameter indicates a large shortest path length between the nodes). Therefore, the attacker can still differentiate two distant nodes. Fig. 7 and Fig. 8 show the diameters and the decay centrality of the aforementioned four graphs for comparison.

We also investigated the differential privacy and prediction uncertainty performance of small-world graphs. Particularly, the impact of edge rewiring probability on Watts-Stongatz small-world graphs [49] is examined. To generate a Watts-Stongatz graph, a ring graph with 500 nodes is first created. Then, each node in the ring is connected with its 10/20/50 nearest neighbors. Finally, each edge (x, y) in the graph is replaced by another edge (x, z) with a rewiring probability, where z is randomly sampled from the node set V. Considering that regular graphs provide good differential privacy and prediction uncertainty performance, we also consider the impact of edge rewiring on regular graphs. Similarly, we first generate a random regular graph[6] with 500 nodes and a degree of 10/20/50. Then, the same edge rewiring as that for Watts-Stongatz graphs is applied. The results are presented in Fig. 9 and Fig. 10.

As the rewiring probability increases, more irregularity is introduced. For regular graph with rewiring, it can be observed that the differential privacy level ϵ increases and the prediction uncertainty c decrease with the increase of rewiring probability, which supports our conjecture that regular graph provides good privacy performance. For Watts-Stongatz graphs, however, the differential privacy level ϵ decreases as the rewiring probability increases. This is mainly because the ring structure has a large diameter, the diameter quickly decreases as the rewiring probability increases, which leads to a smaller ϵ. For the prediction uncertainty, as we can see in Fig. 6 and Fig. 7, the prediction uncertainty of the ring graph remains almost the same as the number of nodes (and therefore the diameter) increases. It seems that the improvement in diameter cannot fully compensate for the negative impact of the introduced irregularity. As a result, the prediction uncertainty c decreases as the rewiring probability increases.

In summary, given the number of nodes, a graph usually has a smaller ϵ and a larger c (and therefore better differential privacy and prediction uncertainty) if it has a small diameter and all the nodes have similar degrees.

D. Privacy Guarantees of the Private Gossip Algorithm within a Candidate Set

Since both differential privacy and prediction uncertainty consider the worst-case scenario, they may not necessarily be good measures for source anonymity. To further illustrate this idea, the differential privacy and prediction uncertainty of the private gossip algorithm within a candidate set is examined in this subsection. More specifically, the following theorem can be proved.

Theorem 3. *Given a general network G and the observation model described in Section III-C with parameter α, the private gossip algorithm is $(\epsilon, 0)$-differentially private within the candidate set Q, where*

$$\epsilon = \ln\left(\max_{k\notin Q, j\neq i\in Q}\left\{\frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]},\right.\right.$$
$$\left.\left.\frac{(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]}\right\}\right). \tag{15}$$

The prediction uncertainty of the private gossip algorithm within the candidate set Q is given by

$$c = \min_{i\in Q, k\notin Q}\left\{\sum_{j\neq i\in Q}\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i],\right.$$
$$\left.\frac{\sum_{j\neq i\in Q}\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[i,k]}\right\}. \tag{16}$$

Proof: Please see Appendix E. ■

Remark 8. *We note that in (15) (similarly (16)), the first term corresponds to the events for which the attacker observes the activities from the nodes in the candidate set before those from the nodes outside the candidate set. Therefore, it is similar to that in Theorem 2. The second term corresponds to the events for which the attacker first observes the activities from the nodes outside the candidate set.*

Fig. 11 and Fig. 12 show the differential privacy level ϵ and prediction uncertainty c of the private gossip algorithm for the three random graphs within a candidate set, respectively. The average degrees of all the graphs are set to 10. In particular, the candidate set Q contains a randomly selected source node and its 1-hop neighbors. 15,000 Monte Carlo runs are performed

[5] Strictly speaking, the 2-dimensional square grid graph is not regular. However, it is close to a regular graph since most of the nodes have the same degree of 4.

[6] Note that there are usually multiple regular graphs with the same degree, given a fixed number of nodes.

This article has been accepted for publication in IEEE Transactions on Network Science and Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2023.3247626

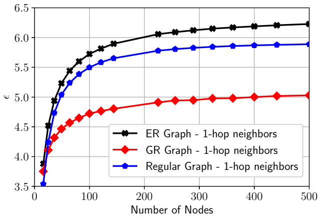


Fig. 11: ϵ v.s. number of nodes ($\alpha = 0.5$)

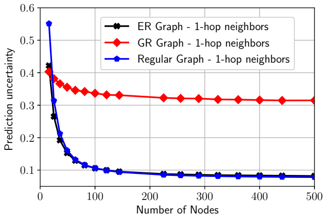


Fig. 12: c v.s. number of nodes ($\alpha = 0.5$)

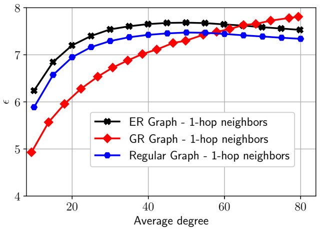


Fig. 13: ϵ v.s. average degree ($n = 500$, $\alpha = 0.5$)

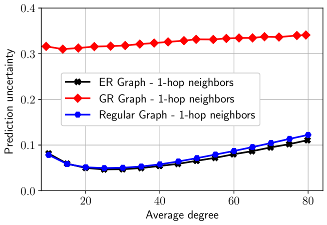


Fig. 14: c v.s. average degree ($n = 500$, $\alpha = 0.5$)

for each graph, and the average ϵ and c within the candidate set Q are presented.

Different from the results in Fig. 5 and Fig. 6, the GR graph performs better than the ER graph and the regular graph. This is because, in the GR graph, there exist some clusters that are well connected locally. In addition, the 1-hop neighboring nodes of a node are likely to be in the same cluster. As a result, it is more difficult for the attacker to distinguish two nodes in the candidate set for the GR graph.

Remark 9. *The above results show that the original differential privacy and prediction uncertainty may have limitations in measuring the privacy guarantees of gossip protocols. Particularly, Fig. 5 and Fig. 6 show that the regular graph has a smaller differential privacy level ϵ and a larger prediction uncertainty c than the GR graph. However, it does not necessarily mean that the regular graph always provides better privacy protection. Fig. 11 and Fig. 12 show that the GR graph performs better than the regular graph in terms of differential privacy and prediction uncertainty within the 1-hop neighbors. In practice, the attacker can often obtain some side information about the source. For instance, the source of the leaked information about a company usually has a close connection to the company. In this case, the source may not be interested to hide itself among all the nodes in the whole network, but instead among those closely related to the company. That being said, if the source node cares more about hiding itself among a subset of nodes (e.g., its 1-hop neighbors), differential privacy and prediction uncertainty within the corresponding candidate set may serve as better privacy metrics.*

Fig. 13 and Fig. 14 show the impact of average degree on ϵ and c for the three random graphs. It can be observed that as the average degree increases, ϵ's of the ER graph and the regular graph first increase and then decrease. Intuitively, ϵ depends on two parameters: the size of the candidate set and the connectivity of the graphs. As the average degree increases, the size of the candidate set becomes larger, while the graphs become better connected. When the average degree is small, the impact of the candidate set size dominates that of the connectivity. As a result, increasing the average degree leads to a larger ϵ. When the average degree is large enough, the impact of the connectivity dominates that of the candidate set size. Therefore, a larger average degree corresponds to a smaller ϵ. The same analysis can be applied to prediction uncertainty. On the other hand, for the GR graph, as the average degree increases, the number of nodes in the candidate set increases. Since the nodes in the candidate set are usually well connected (like a complete graph), it is similar to the complete graph with an increasing number of nodes. As a result, the trends of the curves for the GR graph are similar to those of the complete graph in Fig. 5 and Fig. 6.

V. Privacy-Spreading Tradeoff of Gossip Protocols in Wireless Networks

Up to this point, the communications among the nodes in the network are assumed to be perfect. In many real-world applications, however, the information spreading between two nodes may be realized through wireless communications [3], [50]. With such consideration, the privacy guarantees of gossip protocols in wireless networks are investigated in this section. It is assumed that the communications between the network nodes and between the attacker and its deployed sensors are prone to errors due to various interferences. To simplify the analysis, a failure probability is considered in this setting: Due to interferences, the communications will fail with a probability of f between two nodes during the $gossip$ step, and it is assumed that the attacker fails to receive a report from any of its deployed sensors about the detected events with the same probability f.[7] Note that the failure probability f, induced by detrimental effects in wireless channels, is different from the detection probability α that is due to the limitation in the eavesdropping capability (e.g., computation power) of the sensors. In this case, the privacy guarantees of gossip protocols are characterized by the following theorem.

Theorem 4. *Considering the same setting as in Theorem 1, with the additional constraint that both the legitimate communication and the adversarial reporting fail with a probability f, the gossip-based protocols can guarantee (ϵ, δ)-differential privacy with $\delta \geq \alpha(1-f)$ and c-prediction uncertainty with $c = 0$ in the synchronous setting, and $\delta \geq max[\alpha(1-f) - e^{\epsilon}(1-\alpha(1-f))^{D_G}, \alpha(1-f) - e^{\epsilon}\frac{1-\alpha(1-f)}{n-1}]$ and $c \leq \min_{i \in V} \frac{C_{1-\alpha(1-f)}(i)}{\alpha(1-f)}$ in the asynchronous setting.*

The privacy guarantees of the private gossip algorithm are characterized in the following theorem.

Theorem 5. *Consider the same setting as in Theorem 2, with the additional constraint that both the legitimate communication and the adversarial reporting fail with a probability f, the*

[7] As the first work in this area, this simplified assumption is adopted to facilitate the characterization of the tradeoff between privacy and spreading speed. More realistic assumption concerning two different but correlated failure probabilities [51] warrants further study.

This article has been accepted for publication in IEEE Transactions on Network Science and Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2023.3247626

private gossip algorithm is $(\epsilon, 0)$-differentially private in the asynchronous setting, where

$$\epsilon = \ln\left(\max_{j\neq i\in V}\frac{1}{\alpha(1-f)(I-(1-\alpha(1-f))\hat{A}_i)^{-1}[j,i]}\right). \quad (17)$$

The prediction uncertainty of the private gossip algorithm in the asynchronous setting is given by

$$c = \min_{i\in V}\sum_{j\neq i\in V}\alpha(1-f)(I-(1-\alpha(1-f)\hat{A}_i))^{-1}[j,i]. \quad (18)$$

The proofs of the above theorems readily follow from the previous results and the details are omitted in the interest of space.

Adding artificial noise is a typical way to enhance privacy in practical applications [47]. In wireless networks, interference is a natural source for privacy enhancement as it hampers the attacker's observations of the network activities, which can be further strengthened through approaches such as friendly jamming [52]. However, the information spreading process is impeded as well in such scenarios. The information spreading time of the standard and the private gossip protocols, in this case, is given below.

Theorem 6. *In a wireless network G in which the communications fail with a probability of f, we have*

1) *In the synchronous setting, the private gossip takes $C_G/(1-f)$ rounds on average to inform all nodes in the network, where C_G is the cover time of a random walk in network G.*
2) *In the asynchronous setting, the private gossip takes $C_G/(1-f)$ time on average, while the standard gossip takes $T_{as}/(1-f)$ time on average to finish spreading, where T_{as} is the spreading time of standard gossip when the communication is perfect.*

Proof: Please see Appendix F. ■

Remark 10. *For standard gossip in the synchronous setting, multiple random walks can exist during the spreading process, which renders the analysis of unreliable spreading challenging in general networks. But we conjecture that a similar result as in the synchronous setting may hold.*

The above results indicate a trade-off between privacy and the spreading speed of gossip protocols, which is further explored through simulations below. In particular, following the existing literature in information spreading (e.g., [53], [54]), ER networks and GR Networks with a total number of $n = 100000$ nodes and average node degree of 10 are considered. Each point in the following figures is obtained through simulations with 5 network instances and 100 Monte Carlo runs for each instance. The average 90% spreading time is considered [50]. The privacy-spreading tradeoffs for ER and GR networks for the standard gossip algorithm in the synchronous and asynchronous settings are shown in Fig. 15 and Fig. 16, respectively. It is assumed that $\alpha = 0.5$ and privacy budget $\epsilon = 1$ without loss of generality. The corresponding privacy lower bounds $\underline{\delta}$ in the x-axis are calculated for the considered ER and GR networks using Theorem 4 given the failure probability f (one-to-one correspondence). Similar results are obtained for the private gossip algorithm and omitted here due to the space constraint.

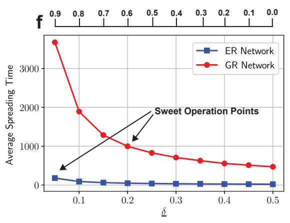


Fig. 15: DP v.s. spreading speed in the synchronous setting

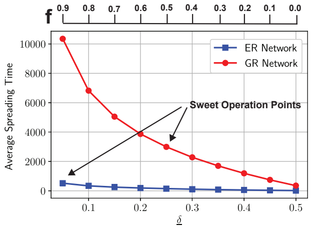


Fig. 16: DP v.s. spreading speed in the asynchronous setting

Remark 11. *Through analysis, it can be seen that the spreading time is inversely proportional to $1-f$ while the privacy lower bound $\underline{\delta}$ is proportional to $1-f$. From Fig. 15 and Fig. 16, it can be seen that when $\underline{\delta}$ increases from 0.05 to 0.1, for GR networks, the average spreading time decreases from around 3600 and 2600 to 1800 and 1300 in the synchronous and asynchronous settings, respectively. This means that we can trade a small loss of privacy for dramatic improvement in spreading time. On the other hand, for ER networks or GR networks with large $\underline{\delta}$ (small f), the average spreading time increases slowly as $\underline{\delta}$ decreases. Therefore, the privacy guarantees of gossip protocols can be strengthened with a small loss of spreading time (e.g., the sweet operation points in Fig. 15 and Fig. 16), which suggests that methods like adding artificial noise can be useful in privacy-preserving information spreading.*

VI. Privacy of Gossip Protocols in Delayed Monitoring

The analyses above assume that the attacker monitors the whole information spreading process right from the beginning. In practice, the attacker may not be able to monitor the ongoing communications in the network as soon as the spreading process starts. Therefore, in this section, we try to quantify the differential privacy of general gossip protocols when the monitoring is delayed. To avoid complications, it is assumed that the communications between nodes and the reception at the attacker are perfect. In addition, the attacker knows the global time in the synchronous setting or the number of communication that has occurred in the asynchronous setting since the beginning of information spreading.

Theorem 7. *Considering the same setting as in Theorem 1, if the attacker starts monitoring the information spreading process t rounds (or t steps of gossip communications in the asynchronous case) after it begins and $t < D_G$, the gossip-based protocols can guarantee (ϵ, δ)-differential privacy with $\delta \geq \frac{1}{d_{max}^t}\alpha$ in the synchronous setting. In the asynchronous setting*

$$\delta \geq \max\left[\frac{1}{d_{max}^t(t+1)!}\alpha - e^{\epsilon}(1-\alpha)^{D_G - t},\right.$$
$$\left.\frac{1}{d_{max}^t(t+1)!}\alpha - e^{\epsilon}\frac{1-\frac{1}{d_{max}^t(t+1)!}\alpha}{n-1}, 0\right], \quad (19)$$

in which $d_{max} = max_{i \in V} d_i$ is the largest node degree.

Proof: Please see Appendix G. ■

Remark 12. *Gossip protocols are not able to protect the source's identity effectively during the early stage of information spreading. As the spreading process continues, more and more randomness is introduced, leading to stronger and stronger privacy. Therefore, in delayed monitoring, it becomes more difficult for the attacker to identify the source node as the delay increases.*

Theorem 8. *Given a general network G and the observation model described in Section III-C with parameter α, if the attacker starts monitoring the information spreading process t steps of gossip communications after it begins, the private gossip algorithm is $(\epsilon, 0)$-differentially private in the asynchronous setting, where*

$$e^{\epsilon} = \max_{i \in V, j \neq z \in V} \frac{A^t[j,i] + \sum_{k \neq i \in V} A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]}{A^t[z,i] + \sum_{k \neq i \in V} A^t[z,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]}. \quad (20)$$

The prediction uncertainty of the private gossip algorithm in the asynchronous setting is given by

$$c = \min_{i \in V, z \in V} \left\{ \frac{\sum_{j \neq i \in V}[A^t[j,z] + \sum_{k \neq z \in V} A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_z)^{-1}[k,z]]}{A^t[i,z] + \sum_{k \neq z \in V} A^t[i,k]\alpha(I-(1-\alpha)\hat{A}_z)^{-1}[k,z]} \right\}. \quad (21)$$

Proof: Please see Appendix H. ■

Remark 13. *Note that in (20), $\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] = P(k \to i) = \alpha\sum_{m=0}^{\infty}(1-\alpha)^m \hat{A}_i^m[k,i]$. For any $k \neq i, \hat{A}_i^0[k,i] = I[k,i] = 0$. Therefore,*

$$\begin{aligned} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] &= \alpha\sum_{m=1}^{\infty}(1-\alpha)^m \hat{A}_i^m[k,i] \\ &\leq \alpha\sum_{m=1}^{\infty}(1-\alpha)^m = 1-\alpha. \end{aligned} \quad (22)$$

As a result,

$$\begin{aligned} &A^t[j,i] + \sum_{k \neq i \in V} A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] \\ &< A^t[j,i] + \sum_{k \neq i \in V} A^t[j,k] = 1. \end{aligned} \quad (23)$$

On the other hand,

$$\begin{aligned} &A^t[z,i] + \sum_{k \neq i \in V} A^t[z,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] \\ &> A^t[z,i] \min_{k \neq i \in V} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] \\ &+ \sum_{k \neq i \in V} A^t[z,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i] \\ &\geq \min_{k \neq i \in V} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]. \end{aligned} \quad (24)$$

Comparing Eq. (20) with Eq. (12), we can see that the ϵ in Theorem 8 is smaller than that in Theorem 2, i.e., in delayed monitoring, the differential privacy guarantee is enhanced. Similar result can be obtained for prediction uncertainty.

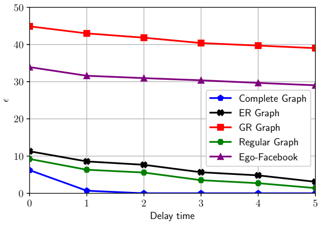


Fig. 17: ϵ v.s. delay time ($\alpha = 0.5$)

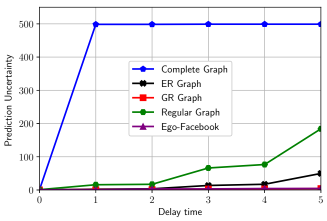


Fig. 18: c v.s. delay time ($\alpha = 0.5$)

Fig. 17 and Fig. 18 show the impact of delay time t on the differential privacy level ϵ and the prediction uncertainty c for different graphs, respectively. In addition to the synthetically generated graphs with 500 nodes, we also examine a real-world network "Ego-Facebook" which consists of friends lists from 4,039 Facebook users [55]. It can be observed that as the delay time increases, ϵ (c) decreases (increases) significantly for the ER graph, the regular graph and the complete graph. This is because, in these graphs, within the delay time t, the information is quickly spread to the rest of the nodes in the graph, which makes it difficult for the attacker to identify the source. For the GR graph and "Ego-Facebook", however, during the delay time, it is likely that the information remains in the same cluster that the source node belongs to. As a result, it is still not difficult for the attacker to distinguish the nodes across clusters.

VII. Conclusions and Future Works

In this work, the privacy guarantees of gossip-based protocols in general networks are investigated. Bounds on differential privacy and prediction uncertainty of gossip protocols in general networks are obtained. For the private gossip algorithm, the differential privacy and prediction uncertainty guarantees are derived in closed form. It is found that these two metrics are closely related to some key network structure parameters, such as network diameter and decay centrality, in the (arguably) more interesting asynchronous setting. Moreover, considering that differential privacy and prediction uncertainty may be restrictive in some scenarios, the relaxed variants of these two metrics are proposed. In wireless networks, through a simplified modeling for unreliable communications, the tradeoff between privacy and spreading efficiency is revealed, and it is suggested that natural or artificial interference can enhance the privacy of gossip protocols with the cost of a decrease in spreading speed. Finally, in delayed monitoring, it is verified that the privacy of gossip protocols is enhanced as the delayed time increases, and the corresponding effect is quantified.

Many interesting problems remain open in this line of research besides those already mentioned above. Investigating more appropriate privacy metrics and designing more effective privacy-aware gossip protocols for different observation models (e.g., network snapshot [8]) and network structures are interesting future directions.

Appendix A
Proof of Theorem 1

Proof: First, for the synchronous setting, let $\mathcal{S}_{i,0}$ be the event that node i's activity is observed by the attacker's sensors at time 0. Then, the probability that such an event happens given the source node is i is $p_G^{(i)}(\mathcal{S}_{i,0}) = \alpha$. If the source node is any other node $j \neq i$, $p_G^{(j)}(\mathcal{S}_{i,0}) = 0$ since node i cannot initialize a communication if it is not a source node at time 0. Therefore, $\delta \geq \alpha$ and $c = 0$.

In the asynchronous setting, let $\mathcal{S}_{i,0}$ be the event that node i's activity is observed by the attacker's sensors as its first observed event. It can be seen that, if the source node is i, then $p_G^{(i)}(\mathcal{S}_{i,0}) = p_G^{(i)}(\mathcal{S}_{i,0}|T_{i,0})p_G^{(i)}(T_{i,0}) + p_G^{(i)}(\mathcal{S}_{i,0}|\overline{T}_{i,0})p_G^{(i)}(\overline{T}_{i,0}) \geq \alpha$, where $T_{i,0}$ stands for the event that the source node is detected during its first communication. If the source node is j, we can consider the following event, denoted as $O_{d(i,j)}$: there is no communication detected by the sensors in the network after $d(i,j)$ gossip actions have been executed from the beginning. Then we have

$$\begin{aligned} p_G^{(j)}(\mathcal{S}_{i,0}) &= p_G^{(j)}(\mathcal{S}_{i,0} \bigcap \overline{O}_{d(i,j)}) + p_G^{(j)}(\mathcal{S}_{i,0} \bigcap O_{d(i,j)}) \\ &= p_G^{(j)}(\mathcal{S}_{i,0} \bigcap O_{d(i,j)}) \\ &\leq p_G^{(j)}(O_{d(i,j)}) = (1-\alpha)^{d(i,j)}, \end{aligned} \tag{25}$$

where the second equality is due to the fact that $\mathcal{S}_{i,0} \bigcap \overline{O}_{d(i,j)} = \emptyset$, as it takes at least $d(i,j)$ communications for the information to be delivered to node i from node j. Since $p_G^{(i)}(\mathcal{S}_{i,0}) \geq \alpha$, by applying Lemma 1, we have

$$\delta \geq \max_{i,j}(\alpha - e^{\epsilon}(1-\alpha)^{d(i,j)}) = \alpha - e^{\epsilon}(1-\alpha)^{D_G}. \tag{26}$$

On the other hand, since $\sum_{j\in V} p_G^{(i)}(\mathcal{S}_{j,0}) = 1$, there exists a node $l \in V$ such that

$$\begin{aligned} p_G^{(i)}(\mathcal{S}_{l,0}) &\leq \frac{1}{n-1}\sum_{j\in V, j\neq i} p_G^{(i)}(\mathcal{S}_{j,0}) \\ &= \frac{1 - p_G^{(i)}(\mathcal{S}_{i,0})}{n-1} \leq \frac{1-\alpha}{n-1}. \end{aligned} \tag{27}$$

This implies $\delta \geq \alpha - e^{\epsilon}\frac{1-\alpha}{n-1}$. By Eq. (26), we have

$$\delta \geq \max\left[\alpha - e^{\epsilon}(1-\alpha)^{D_G}, \alpha - e^{\epsilon}\frac{1-\alpha}{n-1}\right]. \tag{28}$$

Meanwhile, as we have $p_G^{(j)}(\mathcal{S}_{i,0}) \leq (1-\alpha)^{d(i,j)}$, the detection uncertainty can be calculated as

$$\begin{aligned} c = \min_{i,\mathcal{S}}\left(\frac{\sum_{j\neq i} p_G^{(j)}(\mathcal{S})}{p_G^{(i)}(\mathcal{S})}\right) &\leq \min_{i}\frac{\sum_{j\neq i}(1-\alpha)^{d(i,j)}}{\alpha} \\ &= \min_{i\in V}\frac{C_{1-\alpha}(i)}{\alpha}. \end{aligned} \tag{29}$$

■

Appendix B
Proof of Lemma 2

Proof: Let $P_m(j \to i)$ denote the probability that node i is active at time m for the first time before the attacker observes the first event given the source node j. Then we have

$$P(j \to i) = \sum_{m=1}^{\infty} P_m(j \to i). \tag{30}$$

Recall from Remark 5 that the probability of node i being active at time m for the first time is $\hat{A}_i^m[j,i] - \hat{A}_i^{m-1}[j,i]$. Therefore, we have

$$P_m(j \to i) = (1-\alpha)^m[\hat{A}_i^m[j,i] - \hat{A}_i^{m-1}[j,i]]. \tag{31}$$

Plugging (31) into (30) yields

$$\begin{aligned} P(j \to i) &= \sum_{m=1}^{\infty}(1-\alpha)^m[\hat{A}_i^m[j,i] - \hat{A}_i^{m-1}[j,i]] \\ &= \sum_{m=1}^{\infty}(1-\alpha)^m\hat{A}_i^m[j,i] - \sum_{m=1}^{\infty}(1-\alpha)^m\hat{A}_i^{m-1}[j,i] \\ &= \sum_{m=1}^{\infty}(1-\alpha)^m\hat{A}_i^m[j,i] - \sum_{m=1}^{\infty}(1-\alpha)^{m+1}\hat{A}_i^m[j,i] \\ &- (1-\alpha)\hat{A}_i^0[j,i] \\ &= \sum_{m=1}^{\infty}\alpha(1-\alpha)^m\hat{A}_i^m[j,i] - (1-\alpha)I[j,i] \\ &= \alpha\sum_{m=0}^{\infty}(1-\alpha)^m\hat{A}_i^m[j,i]. \end{aligned} \tag{32}$$

Since $\hat{A}_i$ is a right stochastic matrix, it is known that its eigenvalues $|\lambda_{\hat{A}_i}| \leq 1$. Therefore, the eigenvalues of $(1-\alpha)\hat{A}_i$ are strictly less than 1. As a result, (32) can be written alternatively as

$$\begin{aligned} P(j \to i) &= \alpha\sum_{m=0}^{\infty}(1-\alpha)^m\hat{A}_i^m[j,i] \\ &= \alpha(I - (1-\alpha)\hat{A}_i)^{-1}[j,i], \end{aligned} \tag{33}$$

which completes the proof. ■

Appendix C
Proof of Theorem 2

Proof: Let $\mathcal{S}_{i,0}$ denote the event such that node i's activity is observed by the attacker as its first observation. Then, for any $j \neq i$ we have

$$p_G^j(\mathcal{S}_{i,0}) = P(j \to i)p_G^i(\mathcal{S}_{i,0}) \leq p_G^i(\mathcal{S}_{i,0}). \tag{34}$$

Some discussions about (34) are given as follows. Given the source node j, the event $\mathcal{S}_{i,0}$ happens only if: 1) node i becomes active for the first time before the attacker observes its first event (otherwise $\mathcal{S}_{i,0}$ will not happen since node i will not perform *gossip* before it becomes active); 2) the attacker observes node i's activity as its first observation after node i becomes active for the first time. The probability of part 1) is $P(j \to i)$ by its definition. Furthermore, for private gossip, at each time slot, only one node is active. In this case, the probability of part 2) is the same as $p_G^i(\mathcal{S}_{i,0})$.

On the other hand, by Lemma 2

$$\begin{aligned} p_G^i(\mathcal{S}_{i,0}) &= \frac{1}{P(j \to i)}p_G^j(\mathcal{S}_{i,0}) \\ &= \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}p_G^j(\mathcal{S}_{i,0}). \end{aligned} \tag{35}$$

Combining (34) and (35) yields

$$\begin{aligned} p_G^j(\mathcal{S}_{i,0}) \leq p_G^i(\mathcal{S}_{i,0}) &= \frac{1}{P(j \to i)}p_G^j(\mathcal{S}_{i,0}) \\ &\leq \max_{j\neq i\in V}\frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}p_G^j(\mathcal{S}_{i,0}). \end{aligned}$$

As a result, the condition in Definition 1 (i.e., $p_G^i(\mathcal{S}_{i,0}) \leq e^{\epsilon}p_G^j(\mathcal{S}_{i,0})$ and $p_G^j(\mathcal{S}_{i,0}) \leq e^{\epsilon}p_G^i(\mathcal{S}_{i,0})$ hold for any pair of

$i \neq j \in V$) is satisfied for any ϵ such that

$$\begin{aligned} e^{\epsilon} &\geq \max\left\{1, \max_{j\neq i \in V} \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}\right\} \\ &= \max_{j\neq i\in V} \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}. \end{aligned}$$

Since a smaller ϵ indicates better privacy, we have

$$e^{\epsilon} = \max_{j\neq i\in V} \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}. \tag{36}$$

Taking logarithm on both sides of (36) gives the results in Theorem 2.

Regarding the prediction uncertainty, we have

$$\begin{aligned} \frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{i,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{i,0})} &= \frac{\sum_{j\neq i\in V} p_G^{(j)}(\mathcal{S}_{i,0})}{p_G^{(i)}(\mathcal{S}_{i,0})} \\ &= \frac{\sum_{j\neq i\in V} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]p_G^{(i)}(\mathcal{S}_{i,0})}{p_G^{(i)}(\mathcal{S}_{i,0})} \\ &= \sum_{j\neq i\in V} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]. \end{aligned} \tag{37}$$

For any $k \neq i \in V$, it can be shown that

$$\begin{aligned} &\frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{k,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{k,0})} \\ &= \frac{\sum_{j\neq i\in V} p_G^{(j)}(\mathcal{S}_{k,0})}{p_G^{(i)}(\mathcal{S}_{k,0})} = \frac{p_G^{(k)}(\mathcal{S}_{k,0}) + \sum_{j\neq i,k\in V} p_G^{(j)}(\mathcal{S}_{k,0})}{p_G^{(i)}(\mathcal{S}_{k,0})} \\ &= \frac{p_G^{(k)}(\mathcal{S}_{k,0}) + \sum_{j\neq i,k\in V} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]p_G^{(k)}(\mathcal{S}_{k,0})}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]p_G^{(k)}(\mathcal{S}_{k,0})} \\ &= \frac{1+\sum_{j\neq i,k\in V} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]} \\ &\geq \frac{\sum_{j\neq k\in V} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]} \\ &\geq \sum_{j\neq k\in V} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k] = \frac{p_G(I_0 \neq \{k\}|\mathcal{S}_{k,0})}{p_G(I_0 = \{k\}|\mathcal{S}_{k,0})}, \end{aligned} \tag{38}$$

where the first and the second inequality both hold when $P(i \to k) = \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k] = 1$.

As a result.

$$\begin{aligned} c &= \min_{i,\mathcal{S}\subseteq\mathbb{S}} \left(\frac{p_G(I_0 \neq \{i\}|\mathcal{S})}{p_G(I_0 = \{i\}|\mathcal{S})}\right) \\ &= \min_i \sum_{j\neq i\in V} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i], \end{aligned} \tag{39}$$

which completes the proof. ■

Appendix D
Proof of Corollary 1

Proof: The proof of Corollary 1 readily follows from the following results on Gaussian differential privacy.

Theorem. *[56] A mechanism is μ-Gaussian differentially private if and only if it is $(\epsilon, \delta(\epsilon))$-differentially private for all $\epsilon \geq 0$, where*

$$\delta(\epsilon) = \Phi\left(-\frac{\epsilon}{\mu}+\frac{\mu}{2}\right) - e^{\epsilon}\Phi\left(-\frac{\epsilon}{\mu}-\frac{\mu}{2}\right). \tag{40}$$

Given any $(\epsilon, 0)$-differentially private mechanism, it can be shown that it is also μ_1-Gaussian differentially private with $\mu_1 = -2\Phi^{-1}(\frac{1}{1+e^{\epsilon}})$. Applying μ_1 and the above theorem completes the proof. Interested readers may refer to [56] for more details. ■

Appendix E
Proof of Theorem 3

Proof: Let $\mathcal{S}_{i,0}$ denote the event such that node i's activity is observed by the attacker as its first observation. For any $j \neq i \in \mathcal{Q}$, similar to the proof of Theorem 2, we have

$$p_G^{(j)}(\mathcal{S}_{i,0}) = \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]p_G^{(i)}(\mathcal{S}_{i,0}) \leq p_G^{(i)}(\mathcal{S}_{i,0}), \tag{41}$$

and

$$p_G^{(i)}(\mathcal{S}_{i,0}) = \frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]} p_G^{(j)}(\mathcal{S}_{i,0}). \tag{42}$$

For any $k \notin \mathcal{Q}$ and $j \neq i \in \mathcal{Q}$, we have

$$p_G^{(j)}(\mathcal{S}_{k,0}) = \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]p_G^{(k)}(\mathcal{S}_{k,0}), \tag{43}$$

and

$$p_G^{(i)}(\mathcal{S}_{k,0}) = \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]p_G^{(k)}(\mathcal{S}_{k,0}), \tag{44}$$

Therefore,

$$p_G^{(j)}(\mathcal{S}_{k,0}) = \frac{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]} p_G^{(i)}(\mathcal{S}_{k,0}). \tag{45}$$

As a result, following the same analysis as that in the proof of Theorem 2, we have

$$\begin{aligned} e^{\epsilon} = \max_{k\notin\mathcal{Q}, j\neq i\in\mathcal{Q}} \Bigg\{ &\frac{1}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]}, \\ &\frac{(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]} \Bigg\}. \end{aligned} \tag{46}$$

Taking logarithm on both sides of (46) gives the result in Theorem 3.

Regarding the prediction uncertainty, recall that for any $\mathcal{S} \subseteq \mathbb{S}$, we have

$$\frac{p_G(I_0 \neq \{i\}|\mathcal{S})}{p_G(I_0 = \{i\}|\mathcal{S})} = \frac{\sum_{j\neq i\in\mathcal{Q}} p_G^{(j)}(\mathcal{S})}{p_G^{(i)}(\mathcal{S})}. \tag{47}$$

Similar to the proof of Theorem 2, for any event $\mathcal{S}_{i,0} \subseteq \mathbb{S}$ and $j \neq i \in \mathcal{Q}$, it can be shown that

$$\begin{aligned} \frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{i,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{i,0})} &= \frac{\sum_{j\neq i\in\mathcal{Q}} p_G^{(j)}(\mathcal{S}_{i,0})}{p_G^{(i)}(\mathcal{S}_{i,0})} \\ &= \frac{\sum_{j\neq i\in\mathcal{Q}} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]p_G^{(i)}(\mathcal{S}_{i,0})}{p_G^{(i)}(\mathcal{S}_{i,0})} \\ &= \sum_{j\neq i\in\mathcal{Q}} \alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i]. \end{aligned} \tag{48}$$

For any event $\mathcal{S}_{k,0} \subseteq \mathbb{S}$ such that $k \in \mathcal{Q}$,

$$\frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{k,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{k,0})} \geq \frac{p_G(I_0 \neq \{k\}|\mathcal{S}_{k,0})}{p_G(I_0 = \{k\}|\mathcal{S}_{k,0})}. \tag{49}$$

For any event $\mathcal{S}_{k,0} \subseteq \mathbb{S}$ such that $k \notin \mathcal{Q}$, we have

$$\begin{aligned} \frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{k,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{k,0})} &= \frac{\sum_{j\neq i\in\mathcal{Q}} p_G^{(j)}(\mathcal{S}_{k,0})}{p_G^{(i)}(\mathcal{S}_{k,0})} \\ &= \frac{\sum_{j\neq i\in\mathcal{Q}} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]p_G^{(k)}(\mathcal{S}_{k,0})}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]p_G^{(k)}(\mathcal{S}_{k,0})} \\ &= \frac{\sum_{j\neq i\in\mathcal{Q}} \alpha(I-(1-\alpha)\hat{A}_k)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_k)^{-1}[i,k]}. \end{aligned} \tag{50}$$

This article has been accepted for publication in IEEE Transactions on Network Science and Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2023.3247626

Therefore,

$$c=\min_{i\in\mathcal{Q},k\notin\mathcal{Q}}\left\{\sum_{j\neq i\in\mathcal{Q}}\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,i],\ \frac{\sum_{j\neq i\in\mathcal{Q}}\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[j,k]}{\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[i,k]}\right\}, \tag{51}$$

which completes the proof. ■

Appendix F
Proof of Theorem 6

Proof: The spreading process can be considered as a Markov Chain with each state representing the number of informed nodes N_{inf} in the graph. Let $X_k \triangleq \{N_{inf}=k\}$ denote the k-th state of the Markov Chain. For standard gossip in the asynchronous setting and private gossip (in both the synchronous and the asynchronous settings), due to the fact that only one node in the graph is active at any time during the spreading process, the state can only move from X_k to X_{k+1} for all $k\in\{0,1,...,n-1\}$. Each gossip action can be considered as a Bernoulli trial (successful if a previously uninformed node is informed). Given a failure probability of f, the corresponding successful probability decreases by a factor of $1/(1-f)$. In this sense, the expected interstate time, which is essentially the expected time between two consecutive successful gossip actions, is amplified by a factor of $1/(1-f)$. As a result, the final expected time to reach the final state, i.e., the expected spreading time is $1/(1-f)$ times that of the perfect communication scenario. Finally, private gossip is a single random walk in both synchronous and asynchronous settings, and C_G is the expected time to inform all nodes in the graph. ■

Appendix G
Proof of Theorem 7

Proof: In the synchronous setting, consider two nodes i,j such that $d(i,j)=D_G$, and the event that node i's activity is observed by the attacker at the moment when it starts monitoring, which is denoted as $\mathcal{S}_{i,0}$. Considering another node k such that $d(k,i)=t$, the probability that i is informed at round t is

$$p_G^{(k)}(i\in I_t)\geq\prod_{\substack{m\in p_{k\to i}\\ p_{k\to i}:L(p_{k\to i})=t}}\frac{1}{d_m}\geq(\frac{1}{d_{max}})^t, \tag{52}$$

where $p_{k\to i}$ is a path from node k to node i and $L(p_{k\to i})$ is the length of this path. Then $p_G^{(k)}(\mathcal{S}_{i,0})\geq(\frac{1}{d_{max}})^t\alpha$. It is clear that $p_G^{(j)}(\mathcal{S}_{i,0})=0$ since it takes at least $D_G>t$ rounds for the information to be delivered to node i from node j. Therefore, by Lemma 1, $\delta\geq(\frac{1}{d_{max}})^t\alpha$.

In the asynchronous setting, again, consider two nodes i,j such $d(i,j)=D_G$, and let $\mathcal{S}_{i,0}$ denote the event that node i's activity is the first one observed by the attacker. If j is the source node, denote the set of informed and active nodes after t steps of communications as $INA_t(j)$. From this set, find the node $k\in INA_t(j)$ that has the shortest path to node i. Clearly, it requires at least $d(k,i)$ $(\geq(D_G-t))$ steps for the information to reach node i from any node in $INA_t(j)$. Consider $O_{INA_t\to i}$ as the event that no communication is observed by the attacker during the process that the information flows from $INA_t(j)$ to node i. Then,

$$\begin{aligned}p_G^{(j)}(\mathcal{S}_{i,0})&=p_G^{(j)}(\mathcal{S}_{i,0}\bigcap O_{INA_t\to i})\\&\leq p_G^{(j)}(O_{INA_t\to i})\leq(1-\alpha)^{d(k,i)}\leq(1-\alpha)^{D_G-t}.\end{aligned} \tag{53}$$

Also, considering another node l such that $d(l,i)=t$, the probability that node i is informed at the tth step from the beginning of information spreading is

$$p_G^{(l)}(i\in I_t)\geq(\prod_{\substack{m\in p_{l\to i}\\ p_{l\to i}:L(p_{l\to i})=t}}\frac{1}{d_m})\frac{1}{t!}\geq\frac{1}{d_{max}^t t!}, \tag{54}$$

where $\frac{1}{t!}$ is the probability that all nodes in a path $p_{l\to i}$ are activated (whose clocks tick) in a fixed order so that the information reaches node i after t steps from node l. Finally, the probability that node i is activated and its gossip action is observed by the attacker is $\frac{\alpha}{t+1}$. Therefore, $p_G^{(l)}(\mathcal{S}_{i,0})\geq\frac{\alpha}{d_{max}^t(t+1)!}$. By Lemma 1 and the same logic as Eq. (27), we have Eq. (20). ■

Appendix H
Proof of Theorem 8

Proof: Let $p_{G,t}^{(i)}(\mathcal{S}_{i,0})$ denote the probability of the attacker observing $\mathcal{S}_{i,0}$ given source node i in the scenario where the attacker starts monitoring the information spreading process t steps of gossip communications after it begins. In this sense, $p_{G,0}^{(j)}(\mathcal{S}_{i,0})$ is the probability of the event $\mathcal{S}_{i,0}$ if the source is node j and the attacker starts monitoring in the beginning of the information spreading process. In the following, the relationship between $p_{G,t}^{(j)}(\mathcal{S}_{i,0})$ and $p_{G,0}^{(j)}(\mathcal{S}_{i,0})$ is derived, after which similar analysis to the proof of Theorem 2 can be applied to obtain the differential privacy level ϵ. Particularly, for any $j\in V$, we have

$$\begin{aligned}&p_{G,t}^{(j)}(\mathcal{S}_{i,0})\\&=\sum_{k\in V}A^t[j,k]p_{G,0}^{(k)}(\mathcal{S}_{i,0})\\&=A^t[j,i]p_{G,0}^{(i)}(\mathcal{S}_{i,0})+\sum_{k\neq i\in V}A^t[j,k]p_{G,0}^{(k)}(\mathcal{S}_{i,0})\\&=A^t[j,i]p_{G,0}^{(i)}(\mathcal{S}_{i,0})\\&+\sum_{k\neq i\in V}A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]p_{G,0}^{(i)}(\mathcal{S}_{i,0}).\end{aligned} \tag{55}$$

Therefore, following the same analysis as that in the proof of Theorem 2, we have

$$\begin{aligned}e^\epsilon&=\max_{i\in V,j\neq z\in V}\frac{p_{G,t}^{(j)}(\mathcal{S}_{i,0})}{p_{G,t}^{(z)}(\mathcal{S}_{i,0})}\\&=\max_{i\in V,j\neq z\in V}\frac{A^t[j,i]+\sum_{k\neq i\in V}A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]}{A^t[z,i]+\sum_{k\neq i\in V}A^t[z,k]\alpha(I-(1-\alpha)\hat{A}_i)^{-1}[k,i]}.\end{aligned} \tag{56}$$

Regarding the prediction uncertainty, recall that for any

$j \in V$ and $\mathcal{S}_{z,0} \subseteq \mathbb{S}$,

$$\begin{aligned} p^{(j)}_{G,t}(\mathcal{S}_{z,0}) &= A^t[j,z]p^{(z)}_{G,0}(\mathcal{S}_{z,0}) \\ &+ \sum_{k \neq z \in V} A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_z)^{-1}[k,z]p^{(z)}_{G,0}(\mathcal{S}_{z,0}). \end{aligned} \tag{57}$$

Then, for any $i, z \in V$, we have

$$\begin{aligned} \frac{p_G(I_0 \neq \{i\}|\mathcal{S}_{z,0})}{p_G(I_0 = \{i\}|\mathcal{S}_{z,0})} &= \frac{\sum_{j \neq i \in V} p^{(j)}_{G,t}(\mathcal{S}_{z,0})}{p^{(i)}_{G,t}(\mathcal{S}_{z,0})} \\ &= \frac{\sum_{j \neq i \in V}[A^t[j,z] + \sum_{k \neq z \in V} A^t[j,k]\alpha(I-(1-\alpha)\hat{A}_z)^{-1}[k,z]]}{A^t[i,z] + \sum_{k \neq z \in V} A^t[i,k]\alpha(I-(1-\alpha)\hat{A}_z)^{-1}[k,z]}, \end{aligned} \tag{58}$$

which completes the proof. ■

REFERENCES

[1] Y. Huang, R. Jin, and H. Dai, "Differential privacy and prediction uncertainty of gossip protocols in general networks," in *Proc. GLOBECOM*, Taipei, Taiwan, 2020, pp. 1–6.

[2] R. Chandra, V. Ramasubramanian, and K. Birman, "Anonymous gossip: Improving multicast reliability in mobile ad-hoc networks," in *Proc. ICDCS*, Mesa, AZ, USA, 2001, pp. 275–283.

[3] A. D. Dimakis, A. D. Sarwate, and M. J. Wainwright, "Geographic gossip: Efficient averaging for sensor networks," *IEEE Trans. Signal Process*, vol. 56, no. 3, pp. 1205–1216, 2008.

[4] A. J. Ganesh, A.-M. Kermarrec, and L. Massoulié, "Peer-to-peer membership management for gossip-based protocols," *IEEE Trans. Comput.*, vol. 52, no. 2, pp. 139–149, 2003.

[5] B. C. Ferreira, V. Fonte, and J. M. C. Silva, "EAGP: An energy-aware gossip protocol for wireless sensor networks," in *Int. Conf. Softw. Telecommun. Comput. Netw*, Split, Croatia, 2020, pp. 1–6.

[6] P. Bianchi and J. Jakubowicz, "Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 391–405, 2013.

[7] Y. Liu, J. Liu, and T. Basar, "Differentially private gossip gradient descent," in *Proc. CDC*, Miami, FL, USA, 2018, pp. 2777–2782.

[8] J. Jiang, S. Wen, S. Yu, Y. Xiang, and W. Zhou, "Identifying propagation sources in networks: State-of-the-art and comparative studies," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 465–481, 2017.

[9] G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath, "Hiding the rumor source," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6679–6713, 2017.

[10] D. Shah and T. Zaman, "Rumors in a network: Who's the culprit?" *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5163–5181, 2011.

[11] ——, "Rumor centrality: a universal source detector," in *Proc. ACM SIGMETRICS/PERFORMANCE*, London, U.K., 2012, pp. 199–210.

[12] W. Dong, W. Zhang, and C. W. Tan, "Rooting out the rumor culprit from suspects," in *Proc. ISIT*, Istanbul, Turkey, 2013, pp. 2671–2675.

[13] W. Luo, W. P. Tay, and M. Leng, "How to identify an infection source with limited observations," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 586–597, 2014.

[14] K. Zhu and L. Ying, "A robust information source estimator with sparse observations," *Computational Social Networks*, vol. 1, no. 1, pp. 1–21, 2014.

[15] ——, "Information source detection in the SIR model: A sample-path-based approach," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 408–421, 2016.

[16] P. C. Pinto, P. Thiran, and M. Vetterli, "Locating the source of diffusion in large-scale networks," *Phys. Rev. Lett.*, vol. 109, no. 6, p. 068702, 2012.

[17] Z. Shen, S. Cao, W.-X. Wang, Z. Di, and H. E. Stanley, "Locating the source of diffusion in complex networks by time-reversal backward spreading," *Phys. Rev. E*, vol. 93, no. 3, p. 032301, 2016.

[18] F. Yang, S. Yang, Y. Peng, Y. Yao, Z. Wang, H. Li, J. Liu, R. Zhang, and C. Li, "Locating the propagation source in complex networks with a direction-induced search based gaussian estimator," *Knowl.-Based Syst.*, vol. 195, p. 105674, 2020.

[19] A. Kumar, V. S. Borkar, and N. Karamchandani, "Temporally agnostic rumor-source detection," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 3, no. 2, pp. 316–329, 2017.

[20] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Lage-Castellanos, and R. Zecchina, "Bayesian inference of epidemics on networks via belief propagation," *Phys. Rev. Lett.*, vol. 112, no. 11, p. 118701, 2014.

[21] B. Chang, E. Chen, F. Zhu, Q. Liu, T. Xu, and Z. Wang, "Maximum a posteriori estimation for information source detection," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 6, pp. 2242–2256, 2020.

[22] G. Fanti, P. Kairouz, S. Oh, and P. Viswanath, "Spy vs. spy: Rumor source obfuscation," in *Proc. ACM SIGMETRICS/Int. Conf. Measure. Model. Comput. Syst.*, 2015, pp. 271–284.

[23] G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath, "Rumor source obfuscation on irregular trees," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 44, no. 1, pp. 153–164, 2016.

[24] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, 2014.

[25] A. Bellet, R. Guerraoui, and H. Hendrikx, "Who started this rumor? quantifying the natural differential privacy guarantees of gossip protocols," in *Proc. DISC*. Inria, 2020.

[26] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, "Four degrees of separation," in *Proc. 4th ACM Conf. Web Sci.*, 2012, pp. 33–42.

[27] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM rev.*, vol. 42, no. 4, pp. 599–653, 2000.

[28] Z. Zhang, H. Wang, C. Wang, and H. Fang, "Modeling epidemics spreading on social contact networks," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 3, pp. 410–419, 2015.

[29] R. Paluch, X. Lu, K. Suchecki, B. K. Szymański, and J. A. Hołyst, "Fast and accurate detection of spread source in large complex networks," *Sci. Rep.*, vol. 8, no. 1, p. 2508, 2018.

[30] P.-D. Yu, C. W. Tan, and H.-L. Fu, "Epidemic source detection in contact tracing networks: Epidemic centrality in graphs and message-passing algorithms," *IEEE J. Sel. Top. Signal Process.*, vol. 16, no. 2, pp. 234–249, 2022.

[31] H. Kesavareddigari, S. Spencer, A. Eryilmaz, and R. Srikant, "Identification and asymptotic localization of rumor sources using the method of types," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 3, pp. 1145–1157, 2020.

[32] C. Shi, Q. Zhang, and T. Chu, "Source estimation in continuous-time diffusion networks via incomplete observation," *Physica A: Statistical Mechanics and its Applications*, vol. 592, p. 126843, 2022.

[33] Y. Shao, L. Chen, Y. Chen, W. Liu, and C. Dai, "Identifying multiple influence sources in social networks based on latent space mapping," *Inf. Sci.*, 2023.

[34] K. Narayan, H. Agarwal, S. Mittal, K. Thakral, S. Kundu, M. Vatsa, and R. Singh, "Desi: Deepfake source identifier for social media," in *Proc. IEEE/CVF Conf. CVPR*, 2022, pp. 2858–2867.

[35] J. Jiang, S. Wen, S. Yu, Y. Xiang, and W. Zhou, "Rumor source identification in social networks with time-varying topology," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 1, pp. 166–179, 2018.

[36] Q. Huang, C. Zhao, X. Zhang, and D. Yi, "Locating the source of spreading in temporal networks," *Phys. A, Stat. Mech. Appl.*, vol. 468, pp. 434–444, 2017.

[37] Z.-L. Hu, Z. Shen, S. Cao, B. Podobnik, H. Yang, W.-X. Wang, and Y.-C. Lai, "Locating multiple diffusion sources in time varying networks from sparse observations," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, 2018.

[38] L. Fan, B. Li, D. Liu, H. Dai, and Y. Ru, "Identifying propagation source in temporal networks based on label propagation," in *Proc. Int. Conf. Pioneering Comput. Scientists, Eng. Educators (ICPCSEE)*. Springer, 2020, pp. 72–88.

[39] Y. Chai, Y. Wang, and L. Zhu, "Information sources estimation in time-varying networks," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2621–2636, 2021.

[40] L. Cheng, X. Li, Z. Han, T. Luo, L. Ma, and P. Zhu, "Path-based multi-sources localization in multiplex networks," *Chaos, Solitons & Fractals*, vol. 159, p. 112139, 2022.

[41] S. Qu, H. Xu, L. Fu, H. Long, X. Wang, G. Chen, and C. Zhou, "Tracing truth and rumor diffusions over mobile social networks: Who are the initiators," *IEEE Trans. Mob. Comput.*, 2021.

[42] M. Waniek, P. Holme, K. Farrahi, R. Emonet, M. Cebrian, and T. Rahwan, "Trading contact tracing efficiency for finding patient zero," *Sci. Rep.*, vol. 12, no. 1, p. 22582, 2022.

[43] M. Z. Racz and J. Richey, "Rumor source detection with multiple observations under adaptive diffusions," *IEEE Trans. Netw. Sci. Eng.*, 2020.

[44] M. Waniek, P. Holme, M. Cebrian, and T. Rahwan, "Social diffusion sources can escape detection," *Iscience*, vol. 25, no. 9, p. 104956, 2022.

[45] C. Yin, J. Xi, R. Sun, and J. Wang, "Location privacy protection based on differential privacy strategy for big data in industrial internet of things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3628–3636, 2017.

[46] J. Wang, S. Liu, and Y. Li, "A review of differential privacy in individual data release," *Int. J. Distrib. Sensor Netw*, vol. 11, no. 10, p. 259682, 2015.

[47] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur*, 2016, pp. 308–318.

[48] M. O. Jackson, *Social and economic networks*. Princeton university press, 2010.

[49] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[50] H. Zhang, Z. Zhang, and H. Dai, "Gossip-based information spreading in mobile networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5918–5928, 2013.

[51] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 316–329, 2006.

[52] J. P. Vilela, M. Bloch, J. Barros, and S. W. McLaughlin, "Wireless secrecy regions with friendly jamming," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 2, pp. 256–266, 2011.

[53] B. Min, S.-H. Gwak, N. Lee, and K.-I. Goh, "Layer-switching cost and optimality in information spreading on multiplex networks," *Sci. Rep.*, vol. 6, p. 21392, 2016.

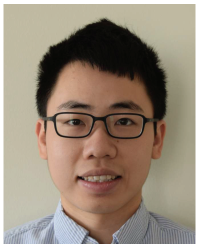
[54] A. Picu, T. Spyropoulos, and T. Hossmann, "An analysis of the information spreading delay in heterogeneous mobility dtns," in *Proc. World Wireless Mobile Multimedia Netw. (WoWMoM)*, 2012, pp. 1–10.

[55] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, Jun. 2014.

[56] J. Dong, A. Roth, and W. J. Su, "Gaussian differential privacy," *arXiv preprint arXiv:1905.02383*, 2019.

Richeng Jin (M'21) received the B.S. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2015, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, NC, USA, in 2020.

He was a Postdoctoral Researcher in electrical and computer engineering at North Carolina State University, Raleigh, NC, USA, from 2021 to 2022. He is currently a faculty member of the department of information and communication engineering with Zhejiang University, Hangzhou, China. His research interests are in the general area of wireless AI, game theory, and security and privacy in machine learning/artificial intelligence and wireless networks.



Yufan Huang received the B.E. degree in electromagnetic engineering from Beihang University, Beijing, China, in 2012, the M.S. degree and the Ph.D. degree in electrical and computer engineering from North Carolina State University, Raleigh, NC, United States, in 2014 and 2019.

His research interests are in the areas of mobile networks, social networks and multi-layer networks. His current research focuses information spreading in multiplex networks.

Zhaoyang Zhang (M'02-SM'21) received his Ph.D. degree from Zhejiang University, Hangzhou, China, in 1998, where he is currently a Qiushi Distinguished Professor. His research interests are mainly focused on the fundamental aspects of wireless communications and networking, such as information theory and coding theory, AI-empowered communications and networking, network signal processing and distributed learning, and synergetic sensing, computing and communication, etc. He has co-authored more than 400 peer-reviewed international journal and conference papers, including 8 conference best papers awarded by IEEE ICC 2019 and IEEE GlobeCom 2020, etc. He was awarded the National Natural Science Fund for Distinguished Young Scholar by NSFC in 2017.

Dr. Zhang is serving or has served as Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS and IET COMMUNICATIONS, etc, and as General Chair, TPC Co-Chair or Symposium Co-Chair for PIMRC 2021 Workshop on Native AI Empowered Wireless Networks, VTC-Spring 2017 Workshop on HMWC, WCSP 2013 / 2018, Globecom 2014 Wireless Communications Symposium, etc. He was also a keynote speaker for Globecom 2021 Workshop on Native-AI Wireless Networks, APCC 2018 and VTC-Fall 2017 Workshop NOMA, etc.

Huaiyu Dai (F'17) received the B.E. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ in 2002.

He was with Bell Labs, Lucent Technologies, Holmdel, NJ, in summer 2000, and with AT&T Labs-Research, Middletown, NJ, in summer 2001. He is currently a Professor of Electrical and Computer Engineering with NC State University, Raleigh, holding the title of University Faculty Scholar. His research interests are in the general areas of communication systems and networks, advanced signal processing for digital communications, communication theory, and information theory. His current research focuses on networked information processing and cross layer design in wireless networks, cognitive radio networks, network security, and associated information-theoretic and computation theoretic analysis.

He has served as an editor of IEEE Transactions on Communications, IEEE Transactions on Signal Processing, and IEEE Transactions on Wireless Communications. Currently he is an Area Editor in charge of wireless communications for IEEE Transactions on Communications, and a member of the Executive Editorial Committee for IEEE Transactions on Wireless Communications. He co-chaired the Signal Processing for Communications Symposium of IEEE Globecom 2013, the Communications Theory Symposium of IEEE ICC 2014, and the Wireless Communications Symposium of IEEE Globecom 2014. He was a co-recipient of best paper awards at 2010 IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS 2010), 2016 IEEE INFOCOM BIGSECURITY Workshop, and 2017 IEEE International Conference on Communications (ICC 2017).