Who Started This Rumor? Quantifying the Natural Differential Privacy of Gossip Protocols

Aurélien Bellet
INRIA, Villeneuve d'Ascq, France
aurelien.bellet@inria.fr

Rachid Guerraoui
EPFL, Lausanne, Switzerland
rachid.guerraoui@epfl.ch

Hadrien Hendrikx
PSL, DIENS, INRIA, Paris, France
hadrien.hendrikx@inria.fr

— Abstract

Gossip protocols (also called rumor spreading or epidemic protocols) are widely used to disseminate information in massive peer-to-peer networks. These protocols are often claimed to guarantee privacy because of the uncertainty they introduce on the node that started the dissemination. But is that claim really true? Can the source of a gossip safely hide in the crowd? This paper examines, for the first time, gossip protocols through a rigorous mathematical framework based on differential privacy to determine the extent to which the source of a gossip can be traceable. Considering the case of a complete graph in which a subset of the nodes are curious, we study a family of gossip protocols parameterized by a "muting" parameter s: nodes stop emitting after each communication with a fixed probability $1 - s$. We first prove that the standard push protocol, corresponding to the case $s = 1$, does not satisfy differential privacy for large graphs. In contrast, the protocol with $s = 0$ (nodes forward only once) achieves optimal privacy guarantees but at the cost of a drastic increase in the spreading time compared to standard push, revealing an interesting tension between privacy and spreading time. Yet, surprisingly, we show that some choices of the muting parameter s lead to protocols that achieve an optimal order of magnitude in both privacy and speed. Privacy guarantees are obtained by showing that only a small fraction of the possible observations by curious nodes have different probabilities when two different nodes start the gossip, since the source node rapidly stops emitting when s is small. The speed is established by analyzing the mean dynamics of the protocol, and leveraging concentration inequalities to bound the deviations from this mean behavior. We also confirm empirically that, with appropriate choices of s, we indeed obtain protocols that are very robust against concrete source location attacks (such as maximum a posteriori estimates) while spreading the information almost as fast as the standard (and non-private) push protocol.

1   Introduction

Gossip protocols (also called *rumor spreading* or *epidemic protocols*), in which participants *randomly* choose a neighbor to communicate with, are both simple and efficient means to exchange information in P2P networks [23, 39, 33, 8]. They are a basic building block to propagate and aggregate information in distributed databases [13, 9] and social networks [14, 27], to model the spread of infectious diseases [29], as well as to train machine learning models on distributed datasets [15, 12, 43, 35].

Some of the information gossiped may be sensitive, and participants sharing it may not want to be identified. This can for instance be the case of whistle-blowers or individuals that would like to exercise their right to freedom of expression in totalitarian regimes. Conversely, it may sometimes be important to locate the source of a (computer or biological) virus, or fake news, in order to prevent it from spreading before too many participants get "infected".

There is a folklore belief that gossip protocols inherently guarantee some form of *source anonymity* because participants cannot know who issued the information in the first place [26]. Similarly, identifying "patient zero" for real-world epidemics is known to be a very hard task. Intuitively indeed, random and local exchanges make identification harder. But to what extent? Although some work has been devoted to the design of source location strategies in specific settings [31, 38, 41], the general anonymity claim has never been studied from a pure *privacy* perspective, that is, independently of the very choice of a source location technique. Depending on the use-case, it may be desirable to have strong privacy guarantees (e.g., in anonymous information dissemination) or, on the contrary, we may hope for weak guarantees, e.g., when trying to identify the source of an epidemic. In both cases, it is crucial to precisely quantify the anonymity level of gossip protocols and study its theoretical limits through a principled approach. This is the challenge we take up in this paper for the classic case of gossip dissemination in a complete network graph.

Our first contribution is an information-theoretic model of anonymity in gossip protocols based on (ϵ, δ)-*differential privacy* (DP) [16, 17]. Originally introduced in the database community, DP is a precise mathematical framework recognized as the gold standard for studying the privacy guarantees of information release protocols. In our proposed model, the information to protect is the source of the gossip, and an adversary tries to locate the source by monitoring the communications (and their relative order) received by a subset of f curious nodes. In a computer network, these curious nodes may have been compromised by a surveillance agency; in our biological example, they could correspond to health professionals who are able to identify whether a given person is infected. Our notion of DP then requires that the probability of any possible observation of the curious nodes is almost the same no matter who is the source, thereby limiting the predictive power of the adversary regardless of its actual source location strategy. A distinctive aspect of our model is that the mechanism that seeks to ensure DP comes only from the *natural* randomness and partial observability of gossip protocols, not from additional perturbation or noise which affects the desired output as generally needed to guarantee DP [18]. We believe our adaptation of DP to the gossip context to be of independent interest. We also complement it with a notion of *prediction uncertainty* which guarantees that even unlikely events do not fully reveal the identity of the source under a uniform prior on the source. This property directly upper bounds the probability of success of any source location attack, including the maximum likelihood estimate.

We use our proposed model to study the privacy guarantees of a generic family of gossip protocols parameterized by a *muting parameter s*: nodes have a fixed probability $1 - s$ to stop emitting after each communication (until they receive the rumor again). In our

■ **Table 1** Summary of results to illustrate the tension between privacy and speed. n is the total number of nodes and f/n is the fraction of curious nodes in the graph. $\delta \in [0, 1]$ quantifies differential privacy guarantees (smaller is better). Spreading time is asymptotic in n.

	Muting param.	δ ensuring $(0, \delta)$-DP	Spreading time
Standard push (minimal privacy, maximal speed)	$s = 1$	1	$\mathcal{O}(\log n)$
Muting after infecting (maximal privacy, minimal speed)	$s = 0$	$\frac{f}{n}$	$\mathcal{O}(n \log n)$
Generic parameterized gossip (privacy vs. speed trade-off)	$0 < s < 1$	$s + (1 - s)\frac{f}{n}$	$\mathcal{O}(\log(n)/s)$

biological parallel, this corresponds to the fact that a person stops infecting other people after some time. The muting parameter captures the ability of the protocol to forget initial conditions, thereby helping to conceal the identity of the source. In the extreme case where $s = 1$, we recover the standard "push" gossip protocol [39], and show that it is inherently *not* differentially private for large graphs. In contrast, we also show that, at the other end of the spectrum, choosing $s = 0$ leads to *optimal privacy guarantees* among all gossip protocols.

More generally, we determine *matching upper and lower bounds* on the privacy guarantees of gossip protocols. Essentially, our upper bounds on privacy are obtained by tightly lower bounding the probability that the source node contacts a curious node before another node does, and upper bounding the probability that this happens for a random node fixed in advance, in a way that holds for all gossip algorithms. Remarkably, despite the fact that the source node always has a non-negligible probability of telling the rumor to a curious node first, our results highlight the fact that setting $s = 0$ leads to strong privacy guarantees in several regimes, including the pure $(\epsilon, 0)$-DP as well as prediction uncertainty.

It turns out that, although achieving optimal privacy guarantees, choosing $s = 0$ leads to a very slow spreading time (*log-linear* in the number of nodes n). This highlights an interesting tension between *privacy* and *spreading time*: the two extreme values for the muting parameter s recover the two extreme points of this trade-off. We then show that more balanced trade-offs can be achieved: appropriate choices of the muting parameter lead to gossip protocols that are *near-optimally private* with a spreading time that is *logarithmic* in the size of the graph. In particular, the trade-off between privacy and speed shows up in the constants but, surprisingly, some choices of the parameter lead to protocols that achieve an optimal order of magnitude for both aspects. Our results on this trade-off are summarized in Table 1: for a proportion f/n of curious nodes, one can see that setting the muting parameter $s = f/n$ achieves almost optimal privacy (up to a factor 2) while being substantially faster than $s = 0$ (optimal up to a factor f/n). Similarly, if one wants to achieve $(0, \delta_0)$-differential privacy with $\delta_0 > 2f/n$, then it is possible to set $s = \delta_0/2$ and obtain a protocol that respects the privacy constraint with spreading time $\mathcal{O}(\log(n)/\delta_0)$. From a technical perspective, these privacy results are obtained by showing that only a small fraction of the possible observations by curious nodes have different probabilities when two different nodes start with the gossip. This requires to precisely evaluate the probability of well-chosen worst-case sequences, which is generally hard as randomness is involved both when nodes decide to stop spreading the rumor (with probability $1 - s$) and when they choose who to communicate with. Our parameterized gossip protocol can be seen as a population protocol [4], and we prove its speed by analyzing its mean dynamics and leveraging concentration inequalities to bound the deviations from the mean dynamics.

We support our theoretical findings by an empirical study of our parameterized gossip protocols. The results show that appropriate choices of *s* lead to protocols that are very robust against classical source location attacks (such as maximum a posteriori estimates) while spreading the information almost as fast as the standard (and non-private) push protocol. Crucially, we observe that our differential privacy guarantees are very well aligned with the ability to withstand attacks that leverage background information, e.g., targeting known activists or people who have been to certain places.

The rest of the paper is organized as follows. We first discuss related work and formally introduce our concept of differential privacy for gossip. Then, we study two extreme cases of our parameterized gossip protocol: the standard push protocol, which we show is not private, and a privacy-optimal but slow protocol. This leads us to investigate how to better control the trade-off between speed and privacy. Finally, we present our empirical study and conclude by discussing open questions.

For pedagogical reasons, we keep our model relatively simple to avoid unnecessary technicalities in the derivation and presentation of our results. For completeness, we discuss the impact of possible extensions (e.g., information observed by the adversary, malicious behavior, termination criterion) in the full version of this paper [6]. Due to space limitations, some detailed proofs are also deferred to the full version.

2   Background and Related Work

2.1   Gossiping

The idea of disseminating information in a distributed system by having each node *push* messages to a randomly chosen neighbor, initially coined the *random phone-call model*, dates back to even before the democratization of the Internet [39]. Such protocols, later called *gossip*, *epidemic* or *rumor spreading*, were for instance applied to ensure the consistency of a replicated database system [13]. They have gained even more importance when argued to model spreading of infectious diseases [29] and information dissemination in social networks [14, 27]. Gossip protocols can also be used to compute aggregate queries on a database distributed across the nodes of a network [34, 9], and have recently become popular in federated machine learning [32] to optimize cost functions over data distributed across a large set of peers [15, 12, 43, 35]. Gossip protocols differ according to their interaction schemes, i.e., *pull* or *push*, sometimes combining both [33, 36, 2].

In this work, we focus on the classical *push* form in the standard case of a *complete* graph with n nodes (labeled from 0 to $n-1$). We now define its key communication primitive. Denoting by I the set of informed nodes, `tell_gossip`(i, I) allows an informed node $i \in I$ to tell the information to another node $j \in \{0, ..., n-1\}$ chosen uniformly at random. `tell_gossip`(i, I) returns j (the node that received the message) and the updated I (the new set of informed nodes that includes j). Equipped with this primitive, we can now formally define the class of gossip protocols that we consider in this paper.

▶ **Definition 1** (Gossip protocols). *A gossip protocol on a complete graph is one that (a) terminates almost surely, (b) ensures that at the end of the execution the set of informed nodes I is equal to $\{0, ..., n-1\}$, and (c) can modify I only through calls to* `tell_gossip`*.*

2.2   Locating the Source of the Gossip

Determining the source of a gossip is an active research topic, especially given the potential applications to epidemics and social networks, see [31] for a recent survey. Existing approaches have focused so far on building *source location attacks* that compute or approximate the

maximum likelihood estimate of the source given some observed information. Each approach typically assumes a specific kind of graphs (e.g., trees, small world, *etc.*), dissemination model and observed information. In *rumor centrality* [41], the gossip communication graph is assumed to be fully observed and the goal is to determine the *center* of this graph to deduce the node that started the gossip. Another line of work studies the setting in which some nodes are *curious sensors* that inform a central entity when they receive a message [38]. Gossiping is assumed to happen at random times and the source node is estimated by comparing the different timings at which information reaches the sensors. The proposed attack is natural in trees but does not generalize to highly connected graphs. The work of [22] focuses on hiding the source instead of locating it. The observed information is a snapshot of who has the rumor at a given time. A specific dissemination protocol is proposed to hide the source but the privacy guarantees only hold for tree graphs.

We emphasize that the privacy guarantees (i.e., the probability not to be detected) that can be derived from the above work only hold under the specific attacks considered therein. Furthermore, all approaches rely on maximum likelihood and hence assume a uniform prior on the probability of each node to be the source. The guarantees thus break if the adversary knows that some of the nodes could not have started the rumor, or if he is aware that the protocol is run twice from the same source.

We note that other problems at the intersection of gossip protocols and privacy have been investigated in previous work, such as preventing unintended recipients from learning the rumor [25], and hiding the initial position of agents in a distributed system [28].

2.3 Differential Privacy

While we borrow ideas from the approaches mentioned above (e.g., we assume that a subset of nodes are curious sensors as in [38]), we aim at studying the fundamental limits of *any* source location attack by measuring the amount of information leaked by a gossip scheme about the identity of the source. For this purpose, a general and robust notion of privacy is required. *Differential privacy* [16, 18] has emerged as a gold standard for it holds independently of any assumption on the model, the computational power, or the background knowledge that the adversary may have. Differentially private protocols have been proposed for numerous problems in the fields of databases, data mining and machine learning: examples include computing aggregate and linear counting queries [18], releasing and estimating graph properties [37, 42], clustering [30], empirical risk minimization [10] and deep learning [1].

In this work, we consider the classic version of differential privacy which involves two parameters $\epsilon, \delta \geq 0$ that quantify the privacy guarantee [17]. More precisely, given any two databases $\mathcal{D}_1$ and $\mathcal{D}_2$ that differ in at most one record, a randomized information release protocol $\mathcal{P}$ is said to guarantee (ϵ, δ)-differential privacy if for any possible output S:

$$p(\mathcal{P}(\mathcal{D}_1) \in S) \leq e^\epsilon p(\mathcal{P}(\mathcal{D}_2) \in S) + \delta, \tag{1}$$

where $p(E)$ denotes the probability of event E. Parameter ϵ places a bound on the ratio of the probability of any output when changing one record of the database, while parameter δ is assumed to be small and allows the bound to be violated with small probability. When $\epsilon = 0$, δ gives a bound on the total variation distance between the output distributions, while $\delta = 0$ is sometimes called "pure" ϵ-differential privacy. DP guarantees hold regardless of the adversary and its background knowledge about the records in the database. In our context, the background information could be the knowledge that the source is among a subset of all nodes. Robustness against such background knowledge is crucial in some applications, for instance when sharing secret information that few people could possibly know or when the

source of an epidemic is known to be among people who visited a certain place. Another key feature of differential privacy is *composability*: if (ϵ, δ)-differential privacy holds for a release protocol, then querying this protocol two times about the same dataset satisfies $(2\epsilon, 2\delta)$-differential privacy. This is important for rumor spreading as it enables to quantify privacy when the source propagates multiple rumors that the adversary can link to the same source (e.g., due to the content of the message). We will see in Section 6 that these properties are essential in practice to achieve robustness to concrete source location attacks.

Existing differentially private protocols typically introduce additional *perturbation* (also called *noise*) to hide critical information [18]. In contrast, an original aspect of our work is that we will solely rely on the *natural* randomness and limited observability brought by gossip protocols to guarantee differential privacy.

3   A Model of Differential Privacy for Gossip Protocols

Our first contribution is a precise mathematical framework for studying the fundamental privacy guarantees of gossip protocols. We formally define the inputs of the gossip protocols introduced in Definition 1, the outputs observed by the adversary during their execution, and the privacy notions we investigate. To ease the exposition, we adopt the terminology of information dissemination, but we sometimes illustrate the ideas in the context of epidemics.

3.1   Inputs and Outputs

As described in Section 2.3, differential privacy is a probabilistic notion that measures the privacy guarantees of a protocol based on the variations of its *output* distribution for a change in its *input*. In this paper, we adapt it to our gossip context. We first formalize the *inputs* and *outputs* of gossip protocols, in the case of a *single piece of information to disseminate* (multiple pieces can be addressed through composition, see Section 2.3). At the beginning of the protocol, a single node, the source, has the information (the gossip, or rumor). This node defines the input of the gossip protocol, and it is the actual "database" that we want to protect. Therefore, in our context, input databases in Equation (1) have only 1 record, which contains the identity of the source (an integer between 0 and $n - 1$). Therefore, all possible input databases differ in at most one record, and differential privacy aims at protecting the content of the database, i.e., which node started the rumor.

We now turn to the outputs of a gossip protocol. The execution of a protocol generates an ordered sequence S_{omni} of pairs (i, j) of calls to `tell_gossip` where $(S_{\mathrm{omni}})_t$ corresponds to the t-th time the `tell_gossip` primitive has been called, i is the node on which `tell_gossip` was used and j the node that was told the information. If several calls to `tell_gossip` happen simultaneously, ties are broken arbitrarily. We assume that the messages are received in the same order that they are sent. This protocol can thus be seen as an epidemic population protocol model [4] in which nodes interact using `tell_gossip`. The sequence S_{omni} corresponds to the output that would be observed by an omniscient entity who could eavesdrop on all communications. It is easy to see that, for any execution, the source can be identified exactly from S_{omni} simply by retrieving $(S_{\mathrm{omni}})_0$.

In this work, we focus on adversaries that monitor a set of *curious nodes* $\mathcal{C}$ of size f, i.e. they observe all communications involving a curious node. This model, previously introduced in [38], is particularly meaningful in large distributed networks: while it is unlikely that an adversary can observe the full state of the network at any given time, compromising or impersonating a subset of the nodes appears more realistic. The number of curious nodes is directly linked with the *release mechanism* of DP: while the final state of the system is

always the same (everyone knows the rumor), the information released depends on which messages were received by the curious nodes during the execution. Formally, the output disclosed to the adversary during the execution of the protocol, i.e., the information he can use to try to identify the source, is a subsequence of S_{omni} as defined below.

▶ **Assumption 2.** *The sequence S observed by the adversary through the (random) execution of the protocol is a (random) subsequence $S = ((S_{\mathrm{omni}})_t | (S_{\mathrm{omni}})_t = (i, j)$ with $j \in \mathcal{C})$, that contains all messages sent to curious nodes. The adversary has access to the relative order of tuples in S, which is the same as in S_{omni}, but not to the index t in S_{omni}.*

It is important to note that the adversary does not know which messages were exchanged between non-curious nodes. In particular, he does not know how many messages were sent in total at a given time. As we focus on complete graphs, knowing which curious node received the rumor gives no information on the source node. For a given output sequence S, we write $S_t = i$ to denote that the t-th `tell_gossip` call in S originates from node i. Omitting the dependence on S, we also denote $t_i(j)$ the time at which node j first receives the message (even for the source) and $t_d(j)$ the time at which j first communicates with a curious node.

The ratio f/n of curious nodes determines the probability of the adversary to gather information (the more curious nodes, the more information leaks). For a fixed f, the adversary only becomes weaker as the network grows bigger. Since we would like to study adversaries with fixed power, unless otherwise noted we make the following assumption.

▶ **Assumption 3.** *The ratio of curious nodes f/n is constant.*

Finally, we emphasize that we do not restrict the computational power of the adversary.

3.2 Privacy Definitions

We now formally introduce our privacy definitions. The first one is a direct application of differential privacy (Equation 1) for the inputs and outputs specified in the previous section. To ease notations, we denote by I_0 the source of the gossip (the set of informed nodes at time 0), and for any given $i \in \{0, ..., n-1\}$, we denote by $p_i(E) = p(E|I_0 = \{i\})$ the probability of event E if node i is the source of the gossip. The protocol is therefore abstracted in this notation. Denoting by $\mathcal{S}$ the set of all possible outputs, we say that a gossip protocol is (ϵ, δ)-differentially private if:

$$p_i(S) \leq e^\epsilon p_j(S) + \delta, \quad \forall S \subset \mathcal{S}, \ \forall i, j \in \{0, ..., n-1\}, \tag{2}$$

where $p(S)$ is the probability that the output belongs to the set S. This formalizes a notion of *source indistinguishability* in the sense that any set of output which is likely enough to happen if node i starts the gossip (say, $p_i(S) \geq \delta$) is almost as likely (up to a e^ϵ multiplicative factor) to be observed by the adversary regardless of the source. Note however that when $\delta > 0$, this definition can be satisfied for protocols that release the identity of the source (this can happen with probability δ). To capture the behavior under unlikely events, we also consider the complementary notion of *c-prediction uncertainty* for $c > 0$, which is satisfied if for a uniform prior $p(I_0)$ on source nodes and any $i \in \{0, ..., n-1\}$:

$$p(I_0 \neq \{i\}|S)/p(I_0 = \{i\}|S) \geq c, \tag{3}$$

for any $S \subset \mathcal{S}$ such that $p_i(S) > 0$. Prediction uncertainty guarantees that no observable output S (however unlikely) can identify a node as the source with large enough probability: it ensures that the probability of success of any source location attack is upper bounded

by $1/(1 + c)$. This holds in particular for the maximum likelihood estimate. Prediction uncertainty does not have the robustness of differential privacy against background knowledge, as it assumes a uniform prior on the source. While it can be shown that $(\epsilon, 0)$-DP with $\epsilon > 0$ implies prediction uncertainty, the converse is not true. Indeed, prediction uncertainty is satisfied as soon as no output identifies any node with enough probability, without necessarily making all pairs of nodes indistinguishable as required by DP. We will see that prediction uncertainty allows to rule out some naive protocols that have nonzero probability of generating sequences which reveal the source with certainty.

Thanks to the symmetry of our problem, we consider without loss of generality that node 0 starts the rumor ($I_0 = \{0\}$) and therefore we will only need to verify Equations (2) and (3) for $i = 0$ and $j = 1$.

4 Extreme Privacy Cases

In this section, we study the fundamental limits of gossip in terms of privacy. To do so, we parameterize gossip protocols by a muting parameter $s \in [0, 1]$, as depicted in Algorithm 1. We thereby capture, within a generic framework, a large family of protocols that fit Definition 1 and work as follows. They maintain a set A of *active nodes* (initialized to the source node) which spread the rumor asynchronously and in parallel: this is modeled by the fact that at each step of the protocol, a randomly selected node $i \in A$ invokes the `tell_gossip` primitive to send the rumor to another node (which in turn becomes active), while i also stays active with probability s. This is illustrated in Figure 1. The muting parameter s can be viewed as a randomized version of *fanout* in [21].[1] Algorithm 1 follows the population protocol model [4], and is also related to the SIS epidemic model [29] but in which the rumor never dies regardless of the value of $s \in [0, 1]$ (there always remain some active nodes). Although we present it from a centralized perspective, we emphasize that Algorithm 1 is asynchronous and can be implemented by having active nodes wake up following a Poisson process.

In the rest of this section, we show that extreme privacy guarantees are obtained for extreme values of the muting parameter s.

4.1 Standard Push has Minimal Privacy

The natural case to study first in our framework is when the muting parameter is set to $s = 1$: this corresponds to the standard push protocol [39] in which nodes always keep emitting after they receive the rumor. Theorem 4 shows that, surprisingly, the privacy guarantees of this protocol become arbitrarily bad as the size of the graph increases (keeping the fraction of curious nodes constant).

▶ **Theorem 4** (Standard push is not differentially private). *If Algorithm 1 with $s = 1$ guarantees (ϵ, δ)-DP for all values of n and constant $\epsilon < \infty$, then $\delta = 1$.*

This result may seem counter-intuitive at first since one could expect that it would be more and more difficult to locate the source when the size of the graph increases. Yet, since the ratio of curious nodes is kept constant, this result comes from the fact that the event $\{t_d(0) \leq t_i(1)\}$ (node 0 communicates with a curious node before node 1 gets the message) becomes more and more likely as n grows, hence preventing any meaningful differential

---

[1] Unlike in the classic fanout, nodes start to gossip again each time they receive a message instead of deactivating permanently.

Algorithm 1 Parameterized Gossip.

Require: n {Number of nodes}, k {Source node}, s {Probability for a node to remain active}
Ensure: $I = \{0, \dots, n-1\}$ {All nodes are informed}
 1: $I \leftarrow \{k\}$, $A \leftarrow \{k\}$
 2: **while** $|I| < n$ **do**
 3:     Sample i uniformly at random from A
 4:     $A \leftarrow A \setminus \{i\}$ with probability $1 - s$
 5:     $j, I \leftarrow \texttt{tell\_gossip}(i, I)$, $A \leftarrow A \cup \{j\}$
 6: **end while**
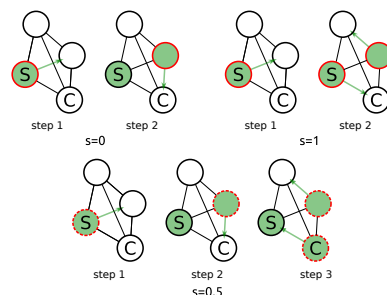
Figure 1 *Left:* Parameterized Gossip. *Right:* Illustration of the role of muting parameter s. S indicates the source and C a curious node. Green nodes know the rumor, and red circled nodes are active. When $s = 0$, there is only one active node at a time, which always stops emitting after telling the gossip. In the case $s = 1$, nodes always remain active once they know the rumor (this is the standard push gossip protocol [39]). When $0 < s < 1$, each node remains active with probability s after each communication.

privacy guarantee when n is large enough. The proof can be found in [6]. Theorem 4 clearly highlights the fact that the standard gossip protocol ($s = 1$) is not differentially private in general. We now turn to the other extreme case, where the muting parameter $s = 0$.

4.2 Muting After Infecting has Maximal Privacy

We now study the privacy guarantees of generic Algorithm 1 when $s = 0$. In this protocol, nodes forward the rumor to exactly one random neighbor when they receive it and then stop emitting until they receive the rumor again. Intuitively, this is good for privacy: the source changes and it is quickly impossible to recover which node started the gossip (as initial conditions are quickly forgotten). In fact, once the source tells the rumor once, the state of the system (the set of active nodes, which in this case is only one node) is completely independent from the source. A similar idea was used in the protocol introduced in [22].

The following result precisely quantifies the privacy guarantees of Algorithm 1 with parameter $s = 0$ and shows that it is *optimally private* among all gossip protocols (in the precise sense of Definition 1).

▶ **Theorem 5.** *Let $\epsilon \geq 0$. For muting parameter $s = 0$, Algorithm 1 satisfies (ϵ, δ)-differential privacy with $\delta = \frac{f}{n} \left(1 - \frac{e^{\epsilon} - 1}{f} \right)$ and c-prediction uncertainty with $c = \frac{n}{f+1} - 1$. Furthermore, these privacy guarantees are optimal among all gossip protocols.*

Proof of Theorem 5. We start by proving the fundamental limits on the privacy of any gossip protocol, and then prove matching guarantees for Algorithm 1 with $s = 0$.

(Fundamental limits in privacy) Proving a lower bound on the differential privacy parameters can be achieved by finding a set of possible outputs S (here, a set of ordered sequences) such that $p_0(S) \geq p_1(S)$. Indeed, a direct application of the definition of Equation (2) yields that given any gossip protocol, $S \subset \mathcal{S}$ and $w_0, w_1 \in \mathbb{R}$ such that $w_0 \leq p_0(S)$ and $p_1(S) \leq w_1$, if the protocol satisfies (ϵ, δ) differential privacy then $\delta \geq w_0 - e^{\epsilon} w_1$.

The proofs need to consider all the messages sent and then distinguish between the ones that are disclosed (sent to curious nodes) and the ones that are not.

Since $I = \{0\}$ then `tell_gossip` is called for the first time by node 0 and it is called at least once otherwise the protocol terminates with $I = \{0\}$, violating the conditions of Definition 1. We denote by $S^{(0)}$ the set of output sequences such that $S_0 = 0$ (i.e., 0 is the first to communicate with a curious node). We also define the event $T_0^c = \{t_d(0) \neq 0\}$ (the source does not send its first message to a curious node). For all $i \notin \mathcal{C} \cup \{0\}$, we have that $p_0(S_0 = i|T_0^c) \leq p_0(S_0 = 0|T_0^c)$ since $p_0(A_1 = \{0\}) = p_0(i \in A_1)$, where A_1 is the set of active nodes at time 1. From this inequality we get

$$\textstyle\sum_{i \notin \mathcal{C}} p_0(S_0 = 0|T_0^c) \geq \sum_{i \notin \mathcal{C}} p_0(S_0 = i|T_0^c) = 1 \geq \sum_{i \notin \mathcal{C}} p_0(S_0 = 1|T_0^c),$$

where the equality comes from the fact that $S_0 = i$ for some $i \notin \mathcal{C}$. The second inequality comes from the fact that $p_j(S_0 = i|T_0^c) = p_j(S_0 = k|T_0^c)$ for all $i, k \neq j$. Therefore, we have $p_0(S_0 = 0|T_0^c) \geq \frac{1}{n-f}$ and $p_0(S_0 = 1|T_0^c) \leq \frac{1}{n-f}$. Combining the above expressions, we derive the probability of $S^{(0)}$ when 0 started the gossip. We write $p_0(S^{(0)}) = p_0(S^{(0)}, t_d(0) = 0) + p_0(S^{(0)}, T_0^c)$ and then, since $p_0(S^{(0)}|t_d(0) = 0) = 1$:

$$p_0(S^{(0)}) = p_0(t_d(0) = 0)p_0(S^{(0)}|t_d(0) = 0) + p_0(S^{(0)}|T_0^c)p_0(T_0^c) \geq \frac{f}{n} + \frac{1}{n-f}\left(1 - \frac{f}{n}\right)$$

In the end, $p_0(S^{(0)}) \geq \frac{f}{n} + \frac{1}{n}$. If node 1 initially has the message, we do the same split and obtain $p_1(S^{(0)}|t_d(0) = 0) = 0$ and so $p_1(S^{(0)}) = p_1(T_0^c)p_1(S^{(0)}|T_0^c) \leq \frac{1}{n}$.

The upper bound on prediction uncertainty is derived using the same quantities:

$$\frac{p(I_0 \neq 0|S^{(0)})}{p(I_0 = 0|S^{(0)})} = \sum_{i \notin C \cup \{0\}} \frac{p_i(S^{(0)})}{p_0(S^{(0)})} \leq (n-f-1)\frac{p_1(S^{(0)})}{p_0(S^{(0)})} \leq \frac{n-f-1}{f+1} = \frac{n}{f+1} - 1.$$

Note that we have never assumed that curious nodes knew how many messages were sent at a given point in time. We have only bounded the probability that the source is the first node that sends a message to curious nodes.

(Matching guarantees for Algorithm 1 with $s = 0$) For this protocol, the only outputs S such that $p_0(S) \neq p_1(S)$ are those in which $t_d(0) = 0$ or $t_d(1) = 0$. We write:

$$p_0(S_0 = 0) = p_0(t_d(0) = 0)p_0(S_0 = 0|t_d(0) = 0) + p_0(T_0^c)p_0(S_0 = 0|T_0^c).$$

For any $i \notin \mathcal{C}$ where $\mathcal{C}$ is the set of curious nodes, we have that $p_0(S_0 = 0|T_0^c) = p_0(S_0 = i|T_0^c) = \frac{1}{n-f}$. Indeed, given that $t_d(0) \neq 0$, the node that receives the first message is selected uniformly at random among non-curious nodes, and has the same probability to disclose the gossip at future rounds. Plugging into the previous equation, we obtain:

$$p_0(S_0 = 0) = \frac{f}{n} + \left(1 - \frac{f}{n}\right)\frac{1}{n-f} = \frac{f+1}{n}.$$

For any other node $i \notin \mathcal{C} \cup \{0\}$, $p_0(S_0 = i) = p_0(T_0^c)p_0(S_0 = i|T_0^c) = \frac{1}{n}$ because $p_0(S_0 = i|t_d(0) = 0) = 0$. Combining these results we get $p_0(S^{(0)}) \leq e^\epsilon p_1(S^{(0)}) + \delta$ for any $\epsilon > 0$ and $\delta = \frac{f}{n}(1 - \frac{e^\epsilon - 1}{f})$. By symmetry, we make a similar derivation for $S^{(1)}$.

To prove the prediction uncertainty result, we use the differential privacy result with $e^\epsilon = f + 1$ (and thus $\delta = 0$) and write that for any $S \in \mathcal{S}$:

$$\frac{p(I_0 \neq 0|S)}{p(I_0 = 0|S)} = \sum_{i \notin C \cup \{0\}} \frac{p_i(S)}{p_0(S)} \geq (n-f-1)e^{-\epsilon} = \frac{n}{f+1} - 1. \qquad \blacktriangleleft$$

Theorem 5 establishes *matching upper and lower bounds* on the privacy guarantees of gossip protocols. More specifically, it shows that setting the muting parameter to $s = 0$ provides strong privacy guarantees that are in fact optimal. Note that in the regime where $\epsilon = 0$ (where DP corresponds to the total variation distance), δ cannot be smaller than the proportion of curious nodes. This is rather intuitive since the source node has probability at least f/n to send its first message to a curious node. However, one can also achieve differential privacy with δ much smaller than f/n by trading-off with $\epsilon > 0$. In particular, the *pure* version of differential privacy ($\delta = 0$) is attained for $\epsilon \approx \log f$, which provides good privacy guarantees when the number of curious nodes is not too large. Furthermore, even though the probability of disclosing *some* information is of order f/n, prediction uncertainty guarantee shows that an adversary with uniform prior always has a high probability of making a mistake when predicting the source. Crucially, these privacy guarantees are made possible by the *natural* randomness and partial observability of gossip protocols.

▶ **Remark 6** (Special behavior of the source). A subtle but key property of Algorithm 1 is that the source follows the same behavior as other nodes. To illustrate how violating this property may give away the source, consider this natural protocol: the source node transmits the rumor to one random node and stops emitting, then standard push (Algorithm 1 with $s = 1$) starts from the node that received the information. While this *delayed start gossip protocol* achieves optimal differential privacy in some regimes, it is fundamentally flawed. In particular, it does not guarantee prediction uncertainty in the sense that $c \to 0$ as the graph grows. Indeed, the adversary can identify the source with high probability by detecting that it communicated only once and then stopped emitting for many rounds. We refer to the full version of this paper [6] for the formal proof.

5 Privacy vs. Speed Trade-offs

While choosing $s = 0$ achieves optimal privacy guarantees, an obvious drawback is that it leads to a very slow protocol since only one node can transmit the rumor at any given time. It is easy to see that the number of gossip operations needed to inform all nodes can be reduced to the time needed for the classical coupon collection problem: it takes $\mathcal{O}(n \log n)$ communications to inform all nodes with probability at least $1 - 1/n$ [19]. As this protocol performs exactly one communication at any given time, it needs time $\mathcal{O}(n \log n)$ to inform all nodes with high probability. This is in stark contrast to the standard push gossip protocol ($s = 1$) studied in Section 4.1 where all informed nodes can transmit the rumor in parallel, requiring only time $\mathcal{O}(\log n)$ [23].

These observations motivate the exploration of the privacy-speed trade-off (with parameter $0 < s < 1$). We first show below that nearly optimal privacy can be achieved for small values of s. Then, we study the spreading time and show that the $\mathcal{O}(\log n)$ time of the standard gossip protocol also holds for $s > 0$, leading to a sweet spot in the privacy-speed trade-off.

5.1 Privacy Guarantees

Theorem 7 conveys a $(0, \delta)$-differential privacy result, which means that apart from some unlikely outputs that may disclose the identity of the source node, most of these outputs actually have the same probability regardless of which node triggered the dissemination. We emphasize that the guarantee we obtain holds for any graph size with fixed proportion f/n of curious nodes.

▶ **Theorem 7** (Privacy guarantees for $s < 1$). *For $0 < s < 1$ and any fixed $r \in \mathbb{N}^*$, Algorithm 1 with muting parameter s guarantees $(0, \delta)$-differential privacy with:*

$$\delta = 1 - (1 - s) \sum_{k=0}^{\infty} s^k \left(1 - \frac{f}{n}\right)^{k+1} \leq 1 - (1 - s^r)\left(1 - \frac{f}{n}\right)^r.$$

For example, choosing $r = 1$ leads to $\delta \leq s + (1 - s)\frac{f}{n}$, as reported in Table 1. Slightly tighter bounds can be obtained, but this is enough already to recover optimal guarantees as $s \to 0$.

Proof. We first consider that S is such that $t_d(0) \geq t_d(1)$. Then, $p_0(S) \leq p_1(S)$ since node 0 needs to receive the rumor before being able to communicate it to curious nodes, and Equation (2) is verified. Suppose now that S is such that $t_d(0) \leq t_d(1)$. In this case, we note t_m the first time at which the source stops to emit (which happens with probability $1 - s$ each time it sends a message). Then, we denote $F = \{t_d(0) \leq t_m\}$ (and F^c its complement). In this case, $p_0(S|F^c) \leq p_1(S|F^c)$. Indeed, conditioned on F^c, $t_d(0) \geq t_i(0)$ if node 0 is not the source and $t_d(0) \geq \max(t_m, t_i(0))$ if it is. Then, we can write:

$$p_0(S) = p_0(S, F^c) + p_0(S, F) \leq p_1(S, F^c) + p_0(F) \leq p_1(S) + p_0(F).$$

Denoting T_f the number of messages after which the source stops emitting, we write:

$$p_0(F) = \sum_{k=1}^{\infty} p_0(T_f = k) p_0(F|T_f = k) = \sum_{k=0}^{\infty} (1 - s) s^k \left(1 - \left(1 - \frac{f}{n}\right)^{k+1}\right), \text{ for } s > 0.$$

Note that we can also write for $k \geq 1$ that $p_0(F) = p_0(F, T_f \leq k) + p_0(F, T_f > k)$, and so:

$$p_0(F) \leq (1 - s^k)\left(1 - \left(1 - \frac{f}{n}\right)^k\right) + s^k = 1 - (1 - s^k)\left(1 - \frac{f}{n}\right)^k. \qquad \blacktriangleleft$$

The differential privacy guarantees given by Theorem 7 and the optimal guarantees of Theorem 5 are of the same order of magnitude when s is of order f/n. Indeed, consider $\epsilon = 0$. Then, setting $r = 1$ in Theorem 7 leads to an additive gap of $s(1 - f/n)$ between the privacy of Algorithm 1 and the optimal guarantee, showing that one can be as close as desired to the optimal privacy as long as s is chosen close enough to 0. In particular, the ratio between the privacy of Algorithm 1 and the lower bound is less than 2 for all $s \leq f/n$. This indicates that the privacy guarantees are very tight in this regime. We also recover exactly the optimal guarantee of Theorem 5 in the case $s = 0$ (without the ability to control the trade-off between ϵ and δ). Importantly, we also show that Algorithm 1 with $s < 1$ satisfies prediction uncertainty, unlike the case where $s = 1$.

▶ **Theorem 8.** *Algorithm 1 guarantees prediction uncertainty with $c = (1 - \frac{f+1}{n})(1 - s)$.*

This result is another evidence that picking $s < 1$ allows to derive meaningful privacy guarantees. The proof can be found in the full version of this paper [6].

5.2 Spreading time

We have shown that parameter s has a significant impact on privacy, from optimal $(s = 0)$ to very weak $(s = 1)$ guarantees. Yet, s also impacts the spreading time: the larger s, the more active nodes at each round. This is highlighted by the two extreme cases, for which the spreading time is already known and exhibits a large gap: $\mathcal{O}(\log n)$ for $s = 1$ and $\mathcal{O}(n \log n)$ for $s = 0$. To establish whether we can obtain a protocol that is both private and fast, we need to characterize the spreading time for the cases where $0 < s < 1$.

The key result of this section is to prove that the logarithmic speed of the standard push gossip protocol holds more generally for all $s > 0$. This result is derived from the fact that the ability to forget does not prevent an *exponential growth* phase. What changes is that the population of active nodes takes approximately $1/s$ rounds to double instead of 1 for standard gossip. For ease of presentation, we state below the result for the synchronous version of Algorithm 1, in which the notion of *round* corresponds to iterating over the full set A. A similar result (with an appropriate notion of rounds) can be obtained for the asynchronous version given in Algorithm 1.

▶ **Theorem 9.** *For a given $s > 0$ and for all $1 > \delta > 0$ and $C \geq 1$, there exists n large enough such that the synchronous version of Algorithm 1 with parameter s sends at least $Cn \log n$ messages in $6C \log(n)/s$ rounds with probability at least $1 - \delta$.*

Proof sketch. The key argument of the proof is that the gossip process very closely follows its mean dynamics. After a transition phase of a logarithmic number of rounds, a constant fraction of the nodes (depending on s) remains active despite the probability to stop emitting after each communication. This "determinism of gossip process" has been introduced in [40], but their analysis only applies to $s = 1$. Our proof accounts for the nontrivial impact of nodes deactivation in the exponential and linear growth phase. Besides, we need to show that in the last phase, with high probability, the population never drops below a critical threshold of active nodes. The full proof is in the full version of this paper [6]. ◀

Theorem 9 shows that generic gossip with $s > 0$ still achieves a logarithmic spreading time even though nodes can stop transmitting the message. The $1/s$ dependence is intuitive since $1/s$ rounds are needed in expectation to double the population of active nodes (without taking collisions into account). Therefore, the exponential growth phase which usually takes time $\mathcal{O}(\log n)$ now takes time $\mathcal{O}(\log(n)/s)$ for $s < 1$. To summarize, we have shown that one can achieve both fast spreading and near-optimal privacy, leading to the values presented in Table 1 of the introduction.

6 Empirical Evaluation

In this section, we evaluate the practical impact of s on the spreading time as well as on the robustness to source location attacks run by adversaries with background knowledge.

6.1 Spreading Time

To complement Theorem 9, which proves logarithmic spreading time (asymptotic in n), we run simulations on a network of size $n = 2^{16}$. The logarithmic spreading time for $s > 0$ is clearly visible in Figure 2a, where we see that the gossip spreads almost as fast for $s = 0.5$ that it does for $s = 1$. We also observe that even when s is small, the gossip remains much faster than for $s = 0$. The results in Figure 2b illustrate that the fraction of active nodes grows exponentially fast for all values of $s > 0$ and then reaches a plateau when the probability of creating a new active node is compensated by the probability of informing an already active node. Empirically, this happens when the fraction of active nodes is of order s.

We note incidentally that gossip protocols are often praised for their robustness to lost messages [3, 24]. While the protocol with $s = 0$ does not tolerate a single lost message, setting $s > 0$ improve the resilience thanks to the linear proportion of active nodes. The latter property makes it unlikely that the protocol stops because of lost messages as long as s is larger than the probability of losing messages. Of course, the protocol remains somewhat sensitive to messages lost during the first few steps.
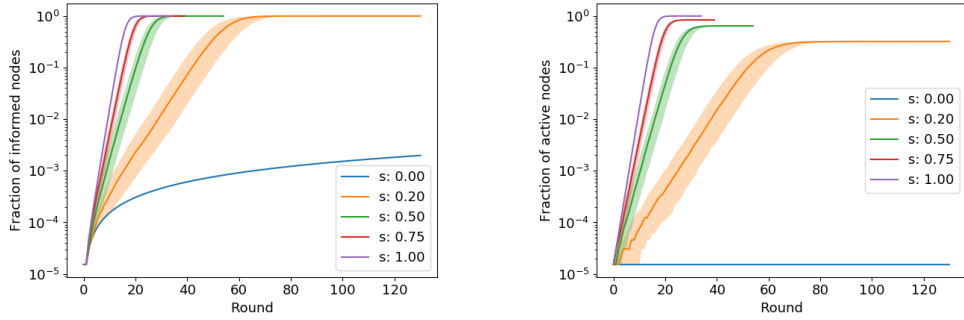
(a) Fraction of informed nodes.

(b) Fraction of active nodes.

■ **Figure 2** Effect of parameter s of Algorithm 1 on the spreading time for a network of $n = 2^{16}$ nodes. The curves represent median values and the shaded area represents the 10 and 90 percent confidence intervals over 100 runs. Each curve stops when all nodes are informed (and so the protocol terminates), except for $s = 0$ since the protocol is very slow in this case.
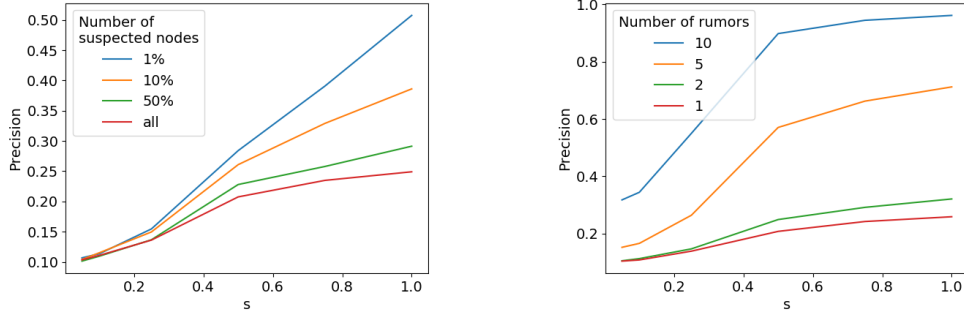
(a) Attack precision under prior information on the source.

(b) Attack precision when the source spreads multiple rumors.

■ **Figure 3** Effect of parameter s of Algorithm 1 on the precision of source location attacks for a network of $n = 2^{16}$ node with 10% of curious nodes. Precision is estimated over 15,000 random runs.

6.2   Robustness Against Source Location Attacks

Getting an intuitive understanding of the privacy guarantees provided by Theorem 7 is not straightforward, as often the case with differential privacy. Therefore, we illustrate the effect of the muting parameter on the guarantees of our gossip protocol by simulating concrete source location attacks. We consider two challenging scenarios where the adversary has some background knowledge: either 1) prior knowledge that the source belongs to a subset of the nodes, or 2) side information indicating that the same source disseminates multiple rumors.

Prior knowledge on the source.   We first consider the case where the adversary is able to narrow down the set of suspected nodes. In this case we can design a provably optimal attack, as shown by the following theorem (see [6]).

▶ **Theorem 10.** *If the adversary has a uniform prior over a subset P of nodes, i.e., $p(I_0 = i) = p(I_0 = j)$ for all $i, j \in P$ and $p(I_0 = i) = 0$ for $i \notin P$, and for some output sequence S, t_c is such that $S_{t_c} \in P$ and $S_t \notin P$ if $t < t_c$, then $p(I_0 = S_{t_c}|S) \geq p(I_0 = i|S)$ for all i.*

Theorem 10 means that under a uniform prior over nodes in P, the attack in which curious nodes predict the source to be the first node in P that communicates with them corresponds to the Maximum A Posteriori (MAP) estimator. The set P represents the prior knowledge of the adversary: he knows for sure that the source belongs to P.

Figure 3a shows the precision of this attack as a function of s for varying degrees of prior knowledge. We see that, when s is small, the prior knowledge does not improve the attack precision significantly, and that the precision remains very close to the probability that the source sends its first message to a curious node. This robustness to prior knowledge is consistent with the properties of differential privacy (see Section 2.3). On the contrary, when s is high (i.e., differential privacy guarantees are weak), the impact of the prior knowledge on the precision of the attack is much stronger.

Multiple dissemination. We investigate another scenario in which differential privacy guarantees can also provide robustness, namely when the adversary knows that the same source node disseminates multiple rumors. In this setting, analytically deriving an optimal attack is very difficult. Instead, we design an attack which leverages the fact that even though the source is not always the first node to communicate with curious nodes, with high probability it will be among the first to do so. More precisely, the curious nodes record the 10 first nodes that communicate with them in each instance (results are not very sensitive to this choice), and they predict the source to be the node that appears in the largest number of instances. In case of a tie, the curious nodes choose the node that first communicated with them, with ties broken at random. Figure 3b shows that the precision of this attack increases dramatically with the number of rumors when s is large, reaching almost sure detection for 10 rumors. Remarkably, for small values of s, the attack precision increases much more gracefully with the number of rumors, as expected from the composition property of differential privacy discussed in Section 2.3. Meaningful privacy guarantees can still be achieved as long as the source does not spread too many rumors.

7 Concluding Remarks

This paper initiates the formal study of privacy in gossip protocols to determine to which extent the source of a gossip can be traceable. Essentially: (1) We propose a formal model of anonymity in gossip protocols based on an adaptation of differential privacy; (2) We establish tight bounds on the privacy of gossip protocols, highlighting their natural privacy guarantees; (3) We precisely capture the trade-off between privacy and speed, showing in particular that it is possible to design both fast and near-optimally private gossip protocols; (4) We experimentally evaluate the speed of our protocols as well as their robustness to source location attacks, validating the relevance of our formal differential privacy guarantees in scenarios where the adversary has some background knowledge.

Our work opens several interesting perspectives. In particular, it paves the way to the study of differential privacy for gossip protocols in *general graphs*, which is challenging and requires relaxations of differential privacy in order to obtain nontrivial guarantees. We refer to the full version of this paper [6] for a discussion of these questions. Another avenue for future research is motivated by very recent work showing that hiding the source of a message can amplify differential privacy guarantees for the *content* of the message [20, 11, 5]. Unfortunately, classic primitives to hide the source of messages such as mixnets can be difficult and costly to deploy. Showing that gossip protocols can *naturally* amplify differential privacy for the message contents would make them very desirable for privacy-preserving distributed AI applications, such as those based on federated [32] and decentralized machine learning [7].

References

1   Martín Abadi, Andy Chu, Ian J. Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *CCS*, 2016.

2   Huseyin Acan, Andrea Collevecchio, Abbas Mehrabian, and Nick Wormald. On the push&pull protocol for rumor spreading. *SIAM Journal on Discrete Mathematics*, 31(2):647–668, 2017.

3   Dan Alistarh, Seth Gilbert, Rachid Guerraoui, and Morteza Zadimoghaddam. How efficient can gossip be?(on the cost of resilient information exchange). In *ICALP*, 2010.

4   Dana Angluin, James Aspnes, and David Eisenstat. Fast computation by population protocols with a leader. *Distributed Computing*, 21(3):183–199, 2008.

5   Borja Balle, James Bell, Adria Gascon, and Kobbi Nissim. The Privacy Blanket of the Shuffle Model. Technical report, arXiv:1903.02837, 2019.

6   Aurélien Bellet, Rachid Guerraoui, and Hadrien Hendrikx. Who started this rumor? Quantifying the natural differential privacy guarantees of gossip protocols. *arXiv preprint arXiv:1902.07138*, 2019.

7   Aurélien Bellet, Rachid Guerraoui, Mahsa Taziki, and Marc Tommasi. Personalized and Private Peer-to-Peer Machine Learning. In *AISTATS*, 2018.

8   Petra Berenbrink, Jurek Czyzowicz, Robert Elsässer, and Leszek Gasieniec. Efficient information exchange in the random phone-call model. *Automata, Languages and Programming*, 2010.

9   Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

10 Kamalika Chaudhuri, Claire Monteleoni, and Anand D. Sarwate. Differentially Private Empirical Risk Minimization. *Journal of Machine Learning Research*, 12:1069–1109, 2011.

11 Albert Cheu, Adam D. Smith, Jonathan Ullman, David Zeber, and Maxim Zhilyaev. Distributed Differential Privacy via Shuffling. Technical report, arXiv:1808.01394, 2018.

12 Igor Colin, Aurélien Bellet, Joseph Salmon, and Stéphan Clémençon. Gossip dual averaging for decentralized optimization of pairwise functions. In *ICML*, 2016.

13 Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. In *PODC*, 1987.

14 Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Social networks spread rumors in sublogarithmic time. In *STOC*, 2011.

15 John C. Duchi, Alekh Agarwal, and Martin J. Wainwright. Dual Averaging for Distributed Optimization: Convergence Analysis and Network Scaling. *IEEE Transactions on Automatic Control*, 57(3):592–606, 2012.

16 Cynthia Dwork. Differential Privacy. In *ICALP*, 2006.

17 Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *EUROCRYPT*, 2006.

18 Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014.

19 Paul Erdős and Alfred Rényi. On a classical problem of probability theory. *Publ. Math. Inst. Hung. Acad. Sci.*, 1961.

20 Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, and Ananth Raghunathan abd Kunal Talwar. Amplification by Shuffling: From Local to Central Differential Privacy via Anonymity. Technical report, arXiv:1811.12469, 2018.

21 Patrick T Eugster, Rachid Guerraoui, Anne-Marie Kermarrec, and Laurent Massoulié. Epidemic information dissemination in distributed systems. *Computer*, 37(5):60–67, 2004.

22 Giulia Fanti, Peter Kairouz, Sewoong Oh, Kannan Ramchandran, and Pramod Viswanath. Hiding the rumor source. *IEEE Transactions on Information Theory*, 2017.

23 Alan M Frieze and Geoffrey R Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1):57–77, 1985.

24 Chryssis Georgiou, Seth Gilbert, Rachid Guerraoui, and Dariusz R Kowalski. Asynchronous gossip. *Journal of the ACM*, 60(2):11, 2013.

25 Chryssis Georgiou, Seth Gilbert, and Dariusz R Kowalski. Confidential gossip. In *ICDCS*, 2011.

26 Mohsen Ghaffari and Calvin Newport. How to discreetly spread a rumor in a crowd. In *DISC*, 2016.

27 George Giakkoupis, Rachid Guerraoui, Arnaud Jégou, Anne-Marie Kermarrec, and Nupur Mittal. Privacy-conscious information diffusion in social networks. In *International Symposium on Distributed Computing*, pages 480–496. Springer, 2015.

28 Karol Gotfryd, Marek Klonowski, and Dominik Pajak. On location hiding in distributed systems. In *International Colloquium on Structural Information and Communication Complexity*, pages 174–192. Springer, 2017.

29 Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

30 Zhiyi Huang and Jinyan Liu. Optimal differentially private algorithms for k-means clustering. In *PODS*, 2018.

31 Jiaojiao Jiang, Sheng Wen, Shui Yu, Yang Xiang, and Wanlei Zhou. Identifying propagation sources in networks: State-of-the-art and comparative studies. *IEEE Communications Surveys & Tutorials*, 19(1):465–481, 2017.

32 Peter Kairouz, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, Rafael G. L. D'Oliveira, Salim El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaid Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara Javidi, Gauri Joshi, Mikhail Khodak, Jakub Konečný, Aleksandra Korolova, Farinaz Koushanfar, Sanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, Rasmus Pagh, Mariana Raykova, Hang Qi, Daniel Ramage, Ramesh Raskar, Dawn Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu, and Sen Zhao. Advances and Open Problems in Federated Learning. Technical report, arXiv:1912.04977, 2019.

33 Richard Karp, Christian Schindelhauer, Scott Shenker, and Berthold Vocking. Randomized rumor spreading. In *FOCS*, 2000.

34 David Kempe, Alin Dobra, and Johannes Gehrke. Gossip-Based Computation of Aggregate Information. In *FOCS*, 2003.

35 Anastasia Koloskova, Sebastian Stich, and Martin Jaggi. Decentralized Stochastic Optimization and Gossip Algorithms with Compressed Communication. In *ICML*, 2019.

36 Dariusz R Kowalski and Christopher Thraves Caro. Estimating time complexity of rumor spreading in ad-hoc networks. In *International Conference on Ad-Hoc Networks and Wireless*, pages 245–256. Springer, 2013.

37 Wentian Lu and Gerome Miklau. Exponential random graph estimation under differential privacy. In *KDD*, 2014.

38 Pedro C Pinto, Patrick Thiran, and Martin Vetterli. Locating the source of diffusion in large-scale networks. *Physical Review Letters*, 2012.

39 Boris Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.

40 Sujay Sanghavi, Bruce Hajek, and Laurent Massoulié. Gossiping with multiple messages. *IEEE Transactions on Information Theory*, 53(12):4640–4654, 2007.

41 Devavrat Shah and Tauhid Zaman. Rumors in a network: Who's the culprit? *IEEE Transactions on Information Theory*, 2011.

42 Haipei Sun, Xiaokui Xiao, Issa Khalil, Yin Yang, Zhan Qin, Hui Wang, and Ting Yu. Analyzing subgraph statistics from extended local views with decentralized differential privacy. In *CCS*, 2019.

43 Paul Vanhaesebrouck, Aurélien Bellet, and Marc Tommasi. Decentralized collaborative learning of personalized models over networks. In *AISTATS*, 2017.