Differentially Private Sublinear Average Degree Approximation

Harry Sivasubramaniam
hsivasub@uwaterloo.ca
University of Waterloo

Haonan Li
h396li@uwaterloo.ca
University of Waterloo

Xi He
xi.he@uwaterloo.ca
University of Waterloo

ABSTRACT

This paper introduces the study of differentially private sublinear average degree approximation for graphs. In the non-private setting, average degree can be approximated within a ratio of $(2+\alpha)$ by sampling $O(\sqrt{n}/\alpha)$ vertices for their degree values. Under ϵ-edge differential privacy, we achieve a ratio of $(2+\alpha+O(\frac{\alpha}{\epsilon\sqrt{n}}))$ with the same sub-linear complexity. For graphs with max degree D, we show a second algorithm that has lower privacy loss with the same noise addition. We also conduct an empirical study of the first estimator on a real world dataset and obtain that the approximation ratio is significantly better in practice.

1 INTRODUCTION

Analysis of graph-like data such as social network and financial transactions plays an important role in gaining business insights and improving services. Two major issues arising in practice are *privacy* and *scalability*. First, such datasets contain sensitive information about individuals and thus require privacy protection. Differential privacy [8, 9] has become an appealing solution and has been extended to graph data analysis [2, 3, 6, 7, 12, 13, 15]. However, these algorithms all require parsing the entire graph, and thus they cannot easily scale to networks with millions or billions of vertices.

In this work, we would like to design sublinear approximation algorithms that satisfy differential privacy for graphs while offering accuracy guarantees. We assume the approximation algorithms can only access the graph using APIs. As each access just queries a vertex and its neighbors, it would be expensive to access the entire graph. We first start with a key parameter of large graphs, the *average degree* of a undirected graph. There is a rich literature on estimating the average degree in the non-private setting [5, 10]. These algorithms access a sublinear number of vertices and aggregate their degree information.

To make these algorithms differentially private, a simple approach is to perturb noise to the final aggregate value with Laplace mechanism [8]. This noise level is proportional to the global sensitivity of the aggregate — the maximum change in the aggregate when adding/removing/changing the sensitive information in the graph. If the queried vertices are correlated, then the global sensitivity can be very large and thus a large amount of noise is required. Hence, we carefully analyzed one of these algorithms, Feige's $(2+\alpha)$-estimator for average degree [10], and designed two algorithms that achieves edge-differential privacy to protect the presence/absence of the edge information in the graph.

Our first algorithm keeps track the frequency of the vertices being queried and provides a tight bound for the global sensitivity of the final estimate of Feige's algorithm. Feige's algorithm samples with replacement the vertices from the graph, and hence some vertices may appear more than once. By taking an appropriate sample size, the frequency of any vertex is bounded by a constant so the global sensitivity is relatively low on most iterations of the algorithm. Our second algorithm directly analyzes the privacy loss through sampling with replacement without keeping track of the frequency of the sampled vertices at a given noise level. Unlike prior work that studies the privacy amplification via subsampling over tabular data [1, 4], the sampling process takes place at vertices that are correlated by edges. In addition, including/excluding a vertex results in different changes in the aggregate in comparison to the absence/presence of an edge for edge-DP. Thus, our second approach does not amplify the privacy to the same extent as the sampling technique over tabular data. We will leave the study of sampling over complex data for the future.

A summary of our contributions is as follows:

- We introduce the study of differentially private sublinear approximation algorithms for graphs.
- We demonstrate average degree can be computed with an approximation factor of $(2+\alpha+O(\frac{\alpha}{\epsilon\sqrt{n}}))$ using $O(\sqrt{n}/\alpha)$ vertex degree queries under ϵ-edge DP.
- We present a privacy analysis for sampling over graphs with max degree D and show it can achieve lower privacy loss than the general case.
- Preliminary empirical study shows that our theoretical work is practical on real data.

Organization. In section 2 we present background information on both query complexity and differential privacy. Section 3 presents our algorithms with accuracy analysis for the degree unbounded case and privacy analysis for both. In section 4 we conduct a preliminary evaluation on the real world dataset and section 5 discusses future directions.

2 BACKGROUND

We describe the background of average degree approximation and differential privacy.

Average Degree Approximation. For a undirected graph $G = (V, E)$ with n vertices, V denotes the set of vertices and E denotes the set of edges in G. For a vertex $v \in V$, let $\deg_G(v)$ denotes the degree of v in G. To efficiently approximate the average degree of a large graph G, $\bar{d}_G = \frac{1}{n}\sum_{v \in V}\deg_G(v)$, Feige [10] proposed an accurate estimator that samples with replacement a sub-linear number of vertices S and queries their degree values. More specifically, the algorithm is described as follows: first, choose a set S by picking at random $s = O(\frac{\sqrt{n}}{\alpha})$ vertices; second, compute the average degree of the vertices in S: $\bar{d}_S = \frac{1}{|S|}\sum_{v \in S}\deg_G(v)$; repeat the above k times and report the minimum of the second step, denoted by $\bar{d}^*$. This estimator has the following accuracy guarantee.

Theorem 1. *Given a graph G with n vertices, Feige's algorithm approximates the average degree with a factor of $(2+\alpha)$ with probability at least $\frac{5}{6}$ by querying the degree value of $O(\frac{\sqrt{n}}{\alpha})$ vertices in total, i.e.,*

$$\Pr[\bar{d}_G < (2+\alpha)\bar{d}^*] \geq 5/6 \text{ and } \Pr[\bar{d}_G > \bar{d}^*] \geq 5/6$$

where $\bar{d}^*$ is the output of Feige's algorithm and $\bar{d}_G$ is the true average degree of G.

We can boost the success probability to at least $1-\delta$ by running the above algorithm $O(\log(\frac{1}{\delta}))$ times and taking the median. The accuracy guarantee can be improved to $(1+\alpha)$ by querying neighboring information of the sampled vertices [11]. As this leaks more information of the graph, this paper first focuses on privatizing Feige's algorithm.

Differential Privacy. Edge differential privacy (DP) [12, 15] is an important variant of differential privacy for releasing statistics of graphs in the literature. This privacy guarantee aims to protect the edge relationship between vertices in the graph. Let $\mathcal{G}_n$ denote the set of n vertex graphs. For any $G, G' \in \mathcal{G}_n$, we use the notation $G \sim G'$ to denote that G and G' differ on exactly one edge.

DEFINITION 1 (Edge Differential Privacy). *A randomized algorithm A is ϵ-edge-DP if for all events R in the output space of A, and for all graphs $G, G' \in \mathcal{G}_n$ with $G \sim G'$ we have* $\Pr[A(G) \in R] \leq e^{\epsilon}\Pr[A(G') \in R]$.

A classic method achieving edge-DP is the Laplace Mechanism [8, 9].

Theorem 2 (Laplace Mechanism). *Let $f : \mathcal{G}_n \to \mathbb{R}$ be any function. The global sensitivity of f is defined as $GS_f = \max_{G \sim G'}|f(G) - f(G')|$. Then the algorithm $A(G) = f(G) + Lap(GS_f/\epsilon)$, satisfies ϵ-edge-DP, where $Lap(x|b) = \frac{1}{2b}e^{-\frac{|x|}{b}}$.*

Algorithm 1 Private Average Degree Estimator

1: **INPUT**: Graph $G = (V, E)$, sampling parameters m, k, s, privacy budget ϵ
2: Initialize $f_{\max} = 0$
3: **for** $i = 1, 2, \ldots, m$ **do**
4:     **for** $j = 1, 2, \ldots, k$ **do**
5:         $S_j \leftarrow$ Sample with replacement s vertices from V
6:         Compute average degree $\tilde{d}_i^j = \sum_{v \in S_j}\deg(v)/s$
7:         $f_{\text{new}} \leftarrow$ Frequency sum of top 2 frequent vertices in S_j
8:         $f_{\max} \leftarrow \max\{f_{\max}, f_{\text{new}}\}$
9:     **end for**
10:     $\tilde{d}_i = \min_j \tilde{d}_i^j$
11: **end for**
12: $\tilde{d} = \text{median}_i(\tilde{d}_i)$
13: $GS = (f_{\max})/s$
14: **return** $\tilde{d} + Lap(GS/\epsilon)$

---

3 PRIVATE AVERAGE DEGREE APPROXIMATION

Given a graph G, we would like to design *a sublinear algorithm to approximate the average degree of G with ϵ-edge DP*. We first present a private estimator that works for general graphs along with the privacy and accuracy analysis. The second subsection is devoted to a similar estimator that works on graphs with maximum degree D. We provide a tighter privacy analysis and obtain a privacy amplification by sub-sampling like result for certain parameter ranges.

3.1 Average Degree for General Graphs

Our algorithm builds upon the average degree estimation proposed by Feige (Section 2) as shown in Algorithm 1. The algorithm takes the minimum of k estimates each obtained by sampling with replacement and then repeats this procedure m times and takes the median (Lines 2-12). We analyzes the global sensitivity of the estimator by tracking the sum of the first and second most frequent sampled vertices over all $m \cdot k$ iterations (Lines 7,8,12). Then we add noise proportional to the global sensitivity using Laplace Mechanism. Next, we show this algorithm satisfies ϵ-edge DP. The sensitivity analysis is non-trivial since we are analysing the sensitivity of the median of minimums of estimates. However, we show that this does not change the global sensitivity value.

Theorem 3. *Algorithm 1 satisfies ϵ-edge-DP.*

Proof. Recall that Algorithm 1 samples from the full set of vertices V, $m \cdot k$ times. We first analyze the sensitivity of the final estimator $\tilde{d}$ when $m = 1$ and $k = 1$. Let S be the sample of size s obtained from sampling with replacement, the algorithm aggregates the degrees for this sample. We would like to see the maximum change to the average degree of S when adding/removing an edge, i.e.,

$$GS_{\tilde{d}} = \max_{G\sim G'} \left| \frac{\sum_{v\in S} \deg_G(v) - \sum_{v\in S} \deg_{G'}(v)}{s} \right| = \frac{f_1(S) + f_2(S)}{s}$$

where $f_i(S)$ is the frequency of the ith most occurring vertex in sample S. Given that G and G' differ on some edge, the global sensitivity corresponds to the case where one endpoint is sampled f_1 times and the other is sampled f_2 times. For any given sample S, $GS_{\tilde{d}}$ is easy to compute as $f_1 + f_2$.

Next, we analyze the case when $k \geq 1$ and $m = 1$. Let the set of samples obtained be $S_1, \ldots, S_k$, where each sample has s vertices. Algorithm 1 computes $\tilde{d}_i$, the average degree of each S_i, and then their minimum value $\tilde{d} = \min_i \tilde{d}_i$. We show that the maximum change to $\tilde{d}$ is still bounded by the $\max_{S_i} f_1(S_i) + f_2(S_i)$ when adding/removing an edge.

Consider a pair of neighboring graphs G, G' where $G' = G + st$ for some edge st. Let $S_a = \text{argmin}_{S_i} \sum_{v\in S_i} \deg_G(v)$ and $S_b = \text{argmin}_{S_i} \sum_{v\in S_i} \deg_{G'}(v)$. Then the difference between estimators before normalizing is

$$|\sum_{v\in S_a} \deg_G(v) - \sum_{v\in S_b} \deg_{G'}(v)| = \sum_{v\in S_b} \deg_{G'}(v) - \sum_{v\in S_a} \deg_G(v)$$

$$\leq \sum_{v\in S_a} (\deg_{G'}(v) - \deg_G(v)) = f_s + f_t \leq f_1(S_a) + f_2(S_a) \leq f_{\max}$$

where f_s and f_t denote the frequency of s and t in S_a.

Last, we analyze when $m \geq 1$ and $k \geq 1$. Similarly consider neighboring graphs G and G'. Fix the set of samples $(S_{1,1}, \ldots, S_{m,k})$, we obtain m minimum values (Line 10). Let $a_1, \ldots, a_m$ be the min values obtained by the algorithm on G and let $b_1, \ldots, b_m$ be the min values obtained on G'. Let $c = \max_{S_i} f_1(S_i) + f_2(S_i) = f_{\max}$ (line 8). Using the lemma 1 below, we get that $GS \leq f_{\max}$ as desired. Since we add noise proportional to this on line 14, by theorem 2 we achieve the desired ϵ-edge DP. □

Lemma 1. *Given sequences $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ where $a_i \leq b_i \leq a_i + c$ for some constant c, then $med(b_i) - med(a_i) \leq c$.*

Next, we show the accuracy analysis of Algorithm 1.

Theorem 4. *For constants m, k, by querying $s \geq \Omega(\frac{\sqrt{n}}{\alpha})$ vertices, Algorithm 1 returns a $(2+\alpha+O(\frac{\alpha}{\epsilon\sqrt{n}}))$-approximation of the average degree with high constant probability.*

Proof. (Sketch) Fix the random vertex samples $S_{1,1}, \ldots, S_{m,k}$, where $S_{i,j}$ are all sampled with replacement from V. From the proof of theorem 3, we know $GS_{\tilde{d}} = \frac{f_{\max}}{s}$. By a balls in bins argument combined with union bound over V and the $m \cdot k$ samples, we can show that $\Pr[f_{\max} \leq r] = 1 - o(1)$ for some $r = O(1)$. Assuming the above holds, by distribution of Laplace noise we get that

$$\Pr[|\hat{d} - \tilde{d}| > t] \leq \exp\left(-\frac{\epsilon t}{GS_{\tilde{d}}}\right) = \exp\left(-\frac{\epsilon t\sqrt{n}}{2r\alpha}\right).$$

Let $t = O(\frac{d\alpha}{\epsilon\sqrt{n}})$, by triangle inequality, $d \leq (2+\alpha+O(\frac{\alpha}{\epsilon\sqrt{n}}))\hat{d}$ with the desired probability. □

3.2 Average Degree for D-Bounded Graphs

Note that the Feige's approximation algorithm is a sampling algorithm. For tabular data, it is well known that sampling helps privacy amplification [1, 4]. However, proving general privacy amplification for edge-DP is non-trivial. When sampling vertices, the resulting structure is a table of vertices, but is not necessarily a meaningful graph. This is unlike the classic privacy amplification results for tabular data as the sub-sampled database is still a valid database in the same domain. This fact is needed when trying to reduce the conditioning of the frequency of an element in the sample. We demonstrate a work around that allows us to show a single estimate gets privacy amplification so long as the graph has max degree D.

Theorem 5. *Given set S of size s is uniformly sampled with replacement from V, the estimator $A_S = \frac{\sum_{i\in S} d_i(G)}{s} + Lap(b)$ satisfies $\left(s \ln\left(\frac{1-2/n+2e^{\frac{D}{sb}}/n}{1-2/n+2e^{-\frac{D}{sb}}/n}\right)\right)$-edge DP.*

Proof. (Sketch) Let $G = (d_1, \ldots, d_n)$ and $G' = (d'_1, \ldots, d'_n)$ be neighbouring degree sequences and without loss of generality assume $|d_{n-1} - d'_{n-1}| = 1$ and $|d_n - d'_n| = 1$. For a set S sampled with replacement, let f_i denote the frequency of element i in S. Also, let S_{s-k} denote the event that $s - k$ sampled vertices are from $\{1, \ldots, n-2\}$ and we let the event $f_{n-1} + f_n$ determine the rest of the behaviour of the sample. Observe that for any output r,

$$\begin{aligned}
& \frac{\Pr_{S,A}[A_S(G) = r | f_{n-1} + f_n = k]}{\Pr_{S,A}[A_S(G) = r | f_{n-1} + f_n = 0]} \\
= & \frac{\sum_{S_{s-k}} \Pr[A_S(G) = r | f_{n-1} + f_n = k, S_{s-k}] \Pr[S_{s-k}]}{\sum_{S_{s-k}} \Pr[A_S(G) = r | f_{n-1} + f_n = 0, S_{s-k}] \Pr[S_{s-k}]} \\
\leq & \max_{S_{s-k}} \frac{\Pr[A_S(G) = r | f_{n-1} + f_n = k, S_{s-k}] \Pr[S_{s-k}]}{\Pr[A_S(G) = r | f_{n-1} + f_n = 0, S_{s-k}] \Pr[S_{s-k}]} \\
= & \max_{S_{s-k}} \frac{\Pr[\eta = r - \frac{\sum_{i\in S} d_i(G)}{s} | f_{n-1} + f_n = k, S_{s-k}]}{\Pr[\eta' = r - \frac{\sum_{i\in S'} d_i(G)}{s} | f_{n-1} + f_n = 0, S_{s-k}]} \leq e^{\frac{kD}{sb}}
\end{aligned}$$

The last inequality above follows as $|\sum_{i\in S} d_i - \sum_{i\in S'} d_i| \leq kD$ where S' differ from S in all first k degrees.

For $k = 0, 1, \ldots, s$, let

$$\begin{aligned}
p_k &= \Pr[f_s + f_t = k] = \binom{s}{k} \left(\frac{2}{n}\right)^k \left(1 - \frac{2}{n}\right)^{s-k} \\
q_k &= \Pr_S[A_S(G) \in R | f_{n-1} + f_n = k] \\
q'_k &= \Pr_S[A_S(G') \in R | f_{n-1} + f_n = k].
\end{aligned}$$

Observe that $q_0 = q'_0$. Note that we have show $q_k \leq e^{\frac{kD}{sb}} q_0$ for each k, we have the following inequality.

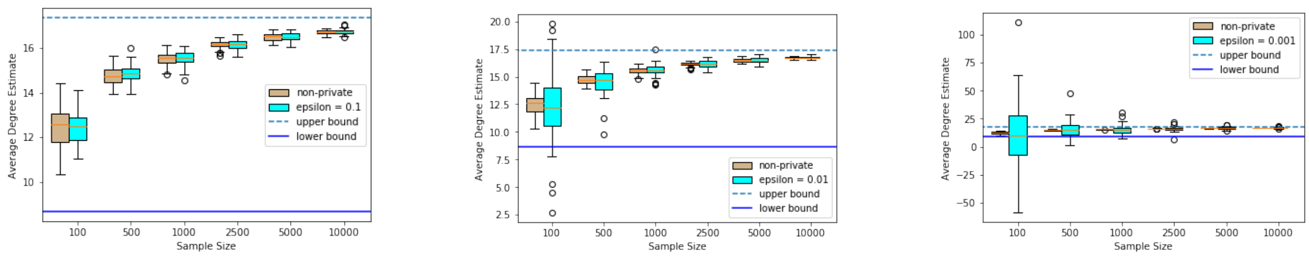

Figure 1: Average degree estimated by Algorithm 1 at different privacy levels. From left to right are results for epsilon values 0.1, 0.01, and 0.001 shown respectively.

$$\Pr_S[A_S(G) \in R] = \sum_{k=0}^{s} p_k q_k \le \sum_{k=0}^{s} \binom{s}{k} \left(\frac{2}{n}\right)^k \left(1 - \frac{2}{n}\right)^{s-k} e^{\frac{kD}{sb}} q_0$$
$$= q_0 \left(1 - 2/n + 2e^{\frac{D}{sb}}/n\right)^s.$$

Similarly, we can show $q'_k \ge e^{\frac{kD}{sb}} q'_0 = e^{\frac{kD}{sb}} q_0$ and so

$$\Pr_S[A_S(G') \in R] \ge q_0 \left(1 - 2/n + 2e^{\frac{-D}{sb}}/n\right)^s.$$

By combining both inequalities above and taking log, we can get the privacy loss of the estimator A_s in the theorem. □

If $\frac{D}{sb} \le 1$, the privacy loss of the estimator A_s is

$$s \ln \left(\frac{1 - 2/n + 2e^{\frac{D}{sb}}/n}{1 - 2/n + 2e^{-\frac{D}{sb}}/n} \right) \le \frac{8D}{bn}.$$

The above estimator and the one from Algorithm 1 both add noise $Lap(b)$ for some b. To achieve ϵ-edge-DP, the above estimator sets $b = \frac{8D}{\epsilon n}$ v.s. Algorithm 1 which sets $b = \frac{f_{\max}}{\epsilon s}$ where $s = O(\frac{\sqrt{n}}{\alpha})$. There are cases where the new analysis for D-bounded graphs yields a much lower privacy loss, up to a $(1/\sqrt{n})$ factor improvement. Another interesting property of this approach is that the noise is not dependent on the frequency of elements in the sample. This leads to a much simpler algorithm. Future work will explore generalizing this approach to handle other estimators and preforming an accuracy analysis similar to that of Algorithm 1.

4 PRELIMINARY RESULTS

We present the evaluation results of our private average degree estimator (Algorithm 1) on real world dataset. This dataset is collected from LiveJournal [14], a free on-line blogging community where users declare friendship each other. The largest connected component of this undirected network is used for our evaluation. It consists of close to 4 million nodes with an average degree of 17.35.

In our experiments, we ensure that Algorithm 1 queries a sublinear number of vertices. We choose $s \in \{100, 500, 1000, 2500, 5000, 10000\}$, $m \in \{1, 3, 5, 10, 15, 20\}$, $k \in \{10\}$, and $\epsilon \in \{0.1, 0.01, 0.001\}$. The default values of s, m, k, ϵ are 10000, 5, 10, and 0.1 respectively. By fixing all of the parameters except one, we observe its effect on the estimation accuracy. We run Algorithm 1 for each setting 50 times and plot the distribution of those 50 iterations on box plots.

Figure 1 shows how the estimates of private and non-private estimators are distributed as s increases, with $m = 5$ and $k = 10$ fixed. The theoretical upper and lower bounds, which are $\hat{d}^*$ (true average degree) and $\hat{d}^*/2$ for this $(2 + \alpha)$-approximation, are also displayed.

We observe that when s increases, the estimate becomes more accurate. Even when the privacy budget is tight, for example $\epsilon = 0.001$, we can still obtain an accurate estimate if s is above $\sqrt{n}$ (2,000 for this dataset). Another interesting observation is that adding some differential privacy noise helps improve the accuracy of the estimates. For example, when $\epsilon = 0.1$ and sample size = 1000, the private estimator sometimes actually gives more accurate results than the non-private estimator. This is because the sampling algorithm underestimates the average degree. When positive noise from the Laplace mechanism is added to the estimate, we can have estimates closer to the true average degree.

Due to space constraint, we leave the evaluation results of our second estimator (Section 3.2) and the parameter study of both algorithms over different datasets to the full paper.

5 DISCUSSION

In this section we discuss some of the future directions for this line of work. First, our current privacy amplification for D bounded graphs has a factor D which can be quite large. We would like to explore removing this dependence on D to get a tighter privacy analysis. It would also be interesting to generalize the analysis to achieve a general privacy amplification lemma for sampling over complex private entities. Average degree is the simplest parameter to study and it is of interest to make private, the sub-linear algorithms for triangle counting and various other subgraph counts. In future work, it will be interesting to study Node-DP as the sensitivities of these estimators is much larger.

REFERENCES

[1] B. Balle, G. Barthe, and M. Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6280–6290, Red Hook, NY, USA, 2018. Curran Associates Inc.

[2] C. Borgs, J. T. Chayes, A. D. Smith, and I. Zadik. Revealing network structure, confidentially: Improved rates for node-private graphon estimation. *CoRR*, abs/1810.02183, 2018.

[3] C. Borgs, J. T. Chayes, A. D. Smith, and I. Zadik. Revealing network structure, confidentially: Improved rates for node-private graphon estimation. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 533–543, 2018.

[4] M. Bun, K. Nissim, U. Stemmer, and S. P. Vadhan. Differentially private release and learning of threshold functions. In V. Guruswami, editor, *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 634–649. IEEE Computer Society, 2015.

[5] A. Dasgupta, R. Kumar, and T. Sarlos. On estimating the average degree. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, page 795–806, New York, NY, USA, 2014. Association for Computing Machinery.

[6] W.-Y. Day, N. Li, and M. Lyu. Publishing graph degree distribution with node differential privacy. In *SIGMOD*, 2016.

[7] X. Ding, X. Zhang, Z. Bao, and H. Jin. Privacy-preserving triangle counting in large graphs. In *CIKM*, 2018.

[8] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *TCC*, 2006.

[9] C. Dwork, A. Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 2014.

[10] U. Feige. On sums of independent random variables with unbounded variance, and estimating the average degree in a graph. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '04, page 594–603, New York, NY, USA, 2004. Association for Computing Machinery.

[11] O. Goldreich and D. Ron. Approximating average parameters of graphs. In J. Díaz, K. Jansen, J. D. P. Rolim, and U. Zwick, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 363–374, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[12] M. Hay, C. Li, G. Miklau, and D. Jensen. Accurate estimation of the degree distribution of private networks. In *ICDM*, 2009.

[13] S. P. Kasiviswanathan, K. Nissim, S. Raskhodnikova, and A. Smith. Analyzing graphs with node differential privacy. In *TCC*, 2013.

[14] LiveJournal. Livejournal social network and ground-truth communities, 2012. data retrieved from LiveJournal, collected by Stanford Network Analysis Project, https://snap.stanford.edu/data/com-LiveJournal.html.

[15] K. Nissim, S. Raskhodnikova, and A. Smith. Smooth sensitivity and sampling in private data analysis. In *STOC*, 2007.