Privacy Preserving Distributed Analysis of Social Networks

Varsha Bhat Kukkala*
Depatment of Computer Science and Engineering
Indian Institute of Technology Ropar
Punjab, India
varsha.bhat@iitrpr.ac.in

ABSTRACT

Social networks have been a popular choice of study, given the surge of online data on friendship networks, communication networks, collaboration networks etc. This popularity, however, is not true for all types of social networks. In the current work, we draw the reader's attention to a class of social networks which are investigated to a limited extent, classified as *distributed sensitive social networks*. It constitutes of networks where the presence or absence of edges in the network is distributedly known to a set of parties, who regard this information as their private data. Supply chain networks, informal networks such as trust network, advice network, enmity network, etc. are a few examples of the same. A major reason for the lack of any substantial study on these networks has been the unavailability of data. As a solution, we propose a privacy preserving approach to investigating these networks. We show the feasibility of using secure multiparty computation techniques to perform the required analysis, while preserving the privacy of every individual's data. The possible approaches that can be considered to ensure the design of efficient secure protocols are discussed such as efficient circuit design, ORAM based secure computation, use of oblivious data structures, etc. The results obtained in the direction of secure network analysis algorithms are also presented.

KEYWORDS

multiparty computation; social networks; distributed networks

1 PROBLEM

Inferring social behavior through analyzing data has been possible with the availability of anonymized datasets of various social networks. Most of the works reported in the literature base their results on the observations made using these datasets. Thus, a majority of the algorithms designed to perform social network analysis consider that the graph $G(V, E)$ representing the anonymized network is available as input. However, it is important to note that the data of certain social networks that capture highly sensitive interactions are not easily available. The focus of this thesis is on scenarios where the social network data is not centrally available, rather is distributedly known to a set of parties. The individual shares are regarded as proprietary data of the parties who do not wish to disclose it for privacy reasons. For example, consider the case of a *supply chain network* (SCN), which is a social network comprising of organizations and captures the buyer-seller interactions among them. An issue in studying these networks is the lack of real world data [23]. The structural position of an organization with respect to the entire SCN plays an important role in determining the profit the organization stands to gain through the network. Hence, trade relations are usually kept private to avoid competitors from taking undue advantage [18]. On the other hand, knowledge of the overall structure of the SCN is beneficial to all the organizations in determining the vulnerabilities and robustness across the global supply chain. This contradicting need to keep the data private as well as be able to use it for analysis has made the study of SCNs challenging.

The conflict between data privacy and its usability is not just seen in the case of SCNs. Several other social networks such as the network of romantic relationships, financial transaction, sexual relationships or the relationship of hatred between employees in an organization, etc., are witnesses to the same. The presence or absence of edges in these networks is regarded private to the corresponding individuals. We refer to such networks as *distributed sensitive social networks* and it is clear that a traditional centralized approach (collection and analysis by a single party) to perform social network analysis is not feasible in these cases. The above claim is well supported by the results of a small scale survey, conducted to determine how private do individuals consider their data [26]. The results of the survey show that the relationship captured in a distributed sensitive social network is regarded more private than other social interactions (see Figure 1). This necessitates the development of a new method to analyze social networks that are both distributed and sensitive in nature. The current work aims at proposing a privacy preserving technique for analyzing distributed sensitive social networks that circumvents the need for disclosing data to a third party.

We model a distributed sensitive social network as a graph $G(V, E)$, where each individual/organization which is a part of the network is depicted by a vertex and V denotes the set of all vertices. An edge $e \in E$ is a pair of vertices (u, v) that captures the private interaction between the vertex u and vertex v. We assume that the set of nodes V is publicly known while the edges are privately known to a set of parties. The problem that we address is for the set of parties to *efficiently* compute network measures on the underlying graph $G(V, E)$, in such a way that none of them disclose their private data, but all of them collectively learn the desired network measure. That is, the set of individuals must engage in running a distributed algorithm to compute the desired network measure,

[*] The author would like to thank her supervisor Dr. Sudarshan Iyengar for the support rendered.

WWW '18 Companion, April 23–27, 2018, Lyon, France

ACM ISBN 978-1-4503-5640-4/18/04.
DOI: https://dx.doi.org/10.1145/3184558.3186578

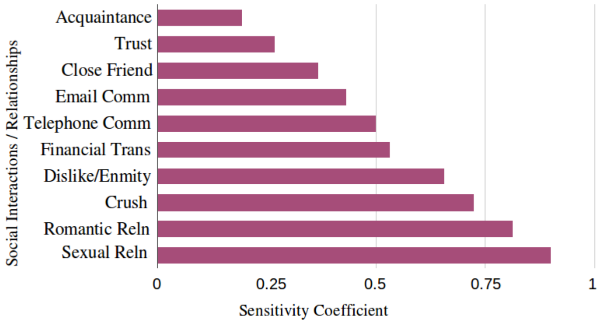


Figure 1: The sensitivity coefficient is defined as the fraction of participants who reported their unwillingness to disclose data to a third party. Therefore, higher the sensitivity, less likely for a centralized approach for data analysis to work. Abbreviations used in the plot: Comm: Communication, Trans: Transaction, Reln: Relationship.

such that the run of the algorithm does not compromise any individual's private data. The algorithm must guarantee *privacy* which ensures that no party learns anything but the output at the end of the algorithm. It must also guarantee *correctness* which ensures that the designed algorithm computes the intended output. This results in computing the desired network measure in a privacy preserving manner and avoids the need for the data holders to outsource the computation to a central authority who learns the complete data, as in the traditional approach.

2 STATE OF THE ART

Despite the privacy concerns associated with distributed sensitive social networks, there are some works that have investigated the structural properties of a few of these networks. It is important to note, however, that all of these works report having carried out the investigation using the traditional approach of a trusted third party model. Some of these works are discussed briefly further.

Investigating the role of the network structure in the spread of STDs has brought the study of sexual networks to the limelight. These networks have been studied under various settings such as internet mediated prostitution [31], online dating [13] and among adolescents in a school [7]. The data for these studies have been collected through third party websites or through archive programs such as Add Health [1] that record data by surveying/interviewing participants. Informal networks are yet another class of networks that can be categorized as distributed sensitive social networks. Any social network that captures the informal interactions of employees in an organization qualifies as an informal network. The network of who trusts whom, who talks to whom, who seeks advice/help from whom, etc. are a few examples for the same. These networks bring out the contrasting nature of the well established formal hierarchy versus the dynamic informal structure. It is well argued that the structure of the informal network determines the effectiveness of the organization [25]. Owing to this, there have been attempts of analyzing many public as well as private companies through questionnaires being posed to its employees [15, 24]. However, a key challenge reported in these studies is eliciting honest answers to the posed questions. The application of social network analysis on supply chains have also been a target of a few studies [14, 18]. One of the drawbacks of these works is that the analysis is performed on small scale supply chain networks due to lack of availability of data [14]. Further, it is well established that the commonly used snowball sampling technique for data collection produces biased samples [35]. In contrast, a socio-centric sampling is preferred, that aims at garnering the complete network. However, getting information of every single dyadic tie in the network has been very challenging. This is all the more difficult when the anonymity of each tie needs to be preserved.

All the previous works involved a third party who learns the private data of the concerned individuals. Although this has the drawback of data not being completely private to the individuals, these works exhibit the importance and benefits of studying the structural properties of the respective networks. As an alternative, there are a few works in the literature that propose privacy preserving approaches to computing a few specific network measures. The problem of securely evaluating influence related metrics over a social network held by a single party like Facebook or Twitter using the activity logs of the individuals in the social network which is present with several service providers like Amazon and Netflix has been studied [33]. Although the question addressed appears to be the same, it is important to note that the setting in which the input data is present is different. The network structure is not distributedly held. Rather, the authors assume that the social network in its entirety is known to a single party while the other parties hold activity logs, each being a (user, action, time) tuple, independent of the network structure. Securely computing the betweenness centrality measure of nodes in a network has been addressed previously in the context of supply chain networks [18] and for criminal networks [22]. Fridgen and Garizy [18] propose a solution in a setting where the notion of security is relaxed, since more than just the centrality score is revealed in the output. The solution by Kerschbaum and Schaad [22] does not incorporate the standard definition of betweenness centrality and has a high computational cost. The solution proposed in the current work addresses both of the drawbacks stated above.

3 PROPOSED APPROACH

The problem addressed in the current thesis can be seen as an instance of *secure multiparty computation* (SMC). This deals with the design of algorithms/protocols that allow a set of n parties (individuals) $P_1, P_2, \ldots P_n$ with private information $x_1, x_2, \ldots x_n$ respectively, to *securely* compute a function f on the aggregate of their inputs. That is, they wish to compute $f(x_1, x_2, \ldots x_n)$. In our case, *parties* refers to the set of individuals who collectively hold the network data. The private input x_i of party P_i is a subset of edges that is privately known to P_i. The function f that the parties are interested in computing is some network measure like that of degree distribution, centrality measures, core-periphery structure, community detection, etc.

The notion of security for an SMC protocol is described using the Real/Ideal paradigm. Given a function f that we are interested in computing, an ideal world refers to the scenario where we assume access to a trusted third party who receives the private inputs from all individuals. He is entrusted to compute f without compromising privacy of individuals nor the integrity of the data. This scenario is known as the *ideal functionality* for f, denoted by $\mathcal{F}_f$. However, in the real world we wish to simulate the same without having access to a third party, by designing a protocol/algorithm (π_f) to compute f. The security of any designed SMC protocol π_f is claimed by proving that for every attacker A on π_f, we can show the existence of an attacker A' on $\mathcal{F}_f$ who would learn/gain the same information/influence as does A. This would show that the designed protocol π_f is in fact as secure as $\mathcal{F}_f$. This notion of security guarantees correctness since the output of the designed protocol π_f and the ideal functionality $\mathcal{F}_f$ are the same. Privacy is also ensured given that each party is guaranteed to learn nothing other than the intended output[1].

SMC has been previously applied to several other applications such as auctions [10], benchmarking [5], computing classical graph algorithms [2, 9] and flow algorithms [3]. However, its applicability for performing social network analysis is less explored. Thus, the aim of the current thesis is to propose a complete suite of protocols designed to perform all the commonly used set of network analysis techniques. This would bridge the gap that currently exists between the use of cryptographic protocols to achieve security and data privacy in social network data mining.

4 METHODOLOGY

Generic constructions have been proposed in the literature that allow the secure evaluation of any computable function. These can be achieved either through garbled circuits [29, 36], boolean circuits [20] or using arithmetic circuits [4, 8, 12]. The complexity of evaluating such a circuit is dependent on the number of gates used (computational and communication complexity) as well as the depth of the circuit (round complexity). These generic constructions can be used in practice only if the functionality f can be represented as a small sized circuit. Thus, the research question boils down to designing efficient circuits specific to each network measure on a given input graph. The focus of the current section will be to highlight the different techniques that can be adopted in order to design efficient solutions.

4.1 Efficient Circuit Design

One approach that can be taken towards designing efficient SMC protocols for computing social network analysis (SNA) measures is to represent the algorithm to compute the SNA measure of interest as a small sized arithmetic circuit consisting of addition and multiplication gates. We can then evaluate such a circuit using the generic constructions for arithmetic circuits. For example, the construction by Damgard et al. [17] realizes the arithmetic ideal functionality $\mathcal{F}_{ABB}$, which provides a generic framework for evaluating reactive tasks expressed as arithmetic circuits. It is designed to be secure in the cryptographic model where the corrupt parties are assumed to be under the complete control of the adversary (*active security*) who is also capable of dynamically varying the set of corrupt parties, as long as only a minority of the parties remain corrupt at a time (*adaptive security*). Further, the performance can be bettered by using the extensions of the $\mathcal{F}_{ABB}$ that provide efficient circuits for other frequently employed operations like comparison, equality check and modulo [16, 30]. That is, the parties can perform these additional operations securely using the primitive operations of the $\mathcal{F}_{ABB}$ ideal functionality itself in constant rounds. Therefore, using any of the state of the art instantiations of the extended $\mathcal{F}_{ABB}$ ideal functionality, we can securely evaluate a circuit designed for a given SNA measure.

4.2 RAM Model Secure Computation

A major source of inefficiency in the circuit based approach to secure computation is the array accesses made over secret indices. That is, consider the scenario where we are interested in accessing the i^{th} index of an array A. If the index i is to be kept private, we would have to construct a circuit that would access all the elements of array A. This results in an $O(n)$ complexity for a single read/write operation, leading to a severe blow-up in cost while performing memory intensive computations, such as in the case of network analysis measures. Thus, RAM based secure computation combines the benefit of ORAM (*oblivious RAM*) schemes with circuit based models of secure computation. An ORAM scheme allows to perform two operations - *initialization* and *access*. The initialization protocol takes the elements of a given array and stores them in a oblivious structure. The access protocol is responsible for translating a given logical address of the array to a sequence of physical addresses of the oblivious structure. The security of an ORAM scheme guarantees that given two arrays of same length, the memory access pattern during the initialization of both arrays are indistinguishable. Similarly, given any two logical addresses, the access protocol generates two indistinguishable sequences of physical memory locations. Thus, an ORAM scheme allows to perform random memory accesses securely in sublinear time. While it may not be obvious on how to achieve an ORAM primitive for memory accesses with efficiency better than $O(n)$, a large amount of research has recently gone into improving the efficiency of this primitive [32, 34].

The use an ORAM scheme in the context of SMC involves the set of parties distributedly holding the shares of the underlying state of the ORAM. The parties then use the circuit based approach to run the initialization and access protocols using the shares of the ORAM state as input [21]. The security of the ORAM scheme ensures that when the sequence of physical memory addresses is revealed to the parties, no private information is leaked and the data retrieved from the accesses is stored in a shared state among the parties. Thus, the retrieved data can be further used as input to the circuit being evaluated. The best known construction of the ORAM primitive for SMC is the Circuit ORAM construction [34], which has asymptotic complexity of $O(\log^2 n)$ to read/write a single entry to an array of size n. However, it has been observed that the scheme

[1] Intermediary data seen during the run of protocol π_f is guaranteed to leak no private information.

provided by Zahur et al. [37] is the most efficient for practical data sizes.

4.3 Oblivious Graph Representation

RAM based secure computation entails to designing *data-oblivious* algorithms, where the primitive operations assumed in the algorithm can be securely evaluated. An algorithm is said to be data-oblivious when the control flow of instructions in the algorithm and the memory accesses made throughout the run are dependent only on the length of the input. That is, given two different inputs of the same length, the behavior of the algorithm remains indistinguishable in both the cases. Thus, when designing data-oblivious algorithms for SNA measures, the representation assumed for storing the network data plays a central role. There are two most widely employed graph data representations, namely, the *adjacency list* and the *adjacency matrix*. Despite being space efficient, the adjacency list representation is never preferred in the data-oblivious setting, since the length of an adjacency list is input dependent (i.e. leaks the degree of the corresponding vertex). On the other hand, the adjacency matrix is well suited for the data-oblivious setting given that it discloses only the number of nodes in the network. Many previous works have employed the adjacency matrix representation for designing data-oblivious algorithms or SMC protocols [9]. However, the adjacency matrix suffers from one major drawback - it is space inefficient. Hence, many problems like single source shortest path and minimum spanning tree that have linear[2] or almost linear running time, the data-oblivious equivalent would have running time $\Omega(|V|^2)$ given the adjacency matrix representation. Furthermore, most real world networks are sparse i.e have $|E| = O(|V|)$ and hence the adjacency matrix representation may not be the best graph representation to opt for. Hence the design of efficient and data-oblivious graph representation is another important research problem that must be addressed for designing efficient data-oblivious algorithms for SNA measures.

4.4 Non-oblivious Secure SMC

A lesser explored approach for designing SMC protocols is to make use of the fact that the output is disclosed in public at the end of an SMC protocol. Hence, one can design algorithms whose control flow and memory access pattern in fact depend on the output of the algorithm. This method is in contrast to the approach of designing data-oblivious algorithms. These protocols, though not data-oblivious, will still give us secure SMC protocols. This methodology also brings out the non-equivalence of data-oblivious algorithms and secure computation protocols.

Brickell and Shmatikov [11] securely compute the single source shortest path, all pair shortest path and the minimum spanning tree of a distributed graph using this approach. The designed protocol reveals the output partially, as and when it is computed. Thus, knowledge of the partial output is used in efficiently computing the rest of the output. To the best of our knowledge, none of the works in the literature have adopted the approach of designing secure MPC protocols that are non-oblivious and harness the fact that the output is leaked in public, in the case of network algorithms.

5 RESULTS

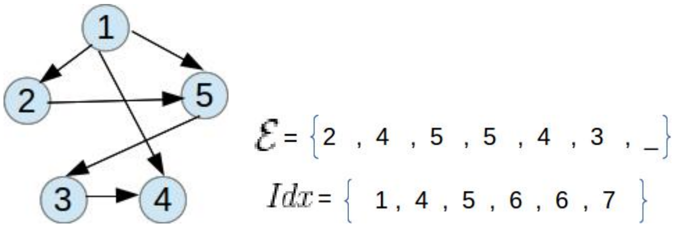
Assuming the adjacency matrix representation of a graph itself, we have designed SMC protocols to compute the degree distribution, closeness centrality, pagerank centrality and the K-shell decomposition of a sensitive network [27]. The construction for degree distribution focuses on computing the in-degree distribution since we assume that each party knows the out-going edges of the corresponding node while is unaware of the in-coming edges. We compute the closeness centrality measure assuming the secure implementation of Dijkstra algorithm provided by Aly and Vyve [3]. The SMC protocol designed to compute pagerank values of nodes in the network is based on the random surfer model [19]. Here, we compute the pagerank values by securely performing a random walk of sufficiently large steps on the underlying graph. The protocol designed for K-shell decomposition is a variation of the standard implementation proposed by Batagelj and Zaversnik [6]. Additionally, we have designed SMC protocol that allows a set of parties to construct an unlabeled isomorphic version of the graph representing their social network [28]. This alternative to constructing the graph itself rather than a specific measure proves useful when we are unsure of the structure and unaware of what measures can help characterize it. All of the above protocols are aimed at designing efficient circuits that can be securely evaluated in the $\mathcal{F}_{ABB}$ hybrid model, as discussed in Section 4.1.

The graph representation assumed while designing protocols also significantly contributes to the complexity. In order to achieve improved efficiency, better suited for sparse real world networks, we propose a data-oblivious graph representation termed the *edgelist* representation, which is space efficient [26]. It uses space linear in the number of nodes and edges in the network. It is also data-oblivious where the representation leaks only the number of nodes and the number of edges in the network and no more information about the network structure is leaked. The edgelist is a $(\mathcal{E}, Idx)$ tuple, where $\mathcal{E}$ is a list consisting of concatenation of ordered set of edges. The list $\mathcal{E}$ begins with all the edges of node 1, followed by edges of node 2, and so on. Idx can be seen as a list of pointers, such that the i^{th} entry of Idx is a pointer to the start location in $\mathcal{E}$ where the edges of node i are present. The details are illustrated through an example as shown in Figure 2. Employing the edgelist graph representation in conjunction with the ORAM primitive (Section 4.2) allowed us to design asymptotically more efficient data-oblivious SNA measures. Using the adjacency matrix representation for designing a data-oblivious variant of the K-shell decomposition algorithm resulted in $O(|V|^3)$ complexity [27]. On the other hand, employing the edgelist representation, we could reduce the complexity of the K-shell decomposition algorithm to $O((|V| + |E|) \log^2 |V|)$ [26].

6 CONCLUSION

In this paper, we introduced the notion of distributed sensitive social networks and highlighted on their unexplored potential. We proposed the use of secure multiparty computation to study the network properties of the distributedly held sensitive networks, while ensuring the privacy of the involved individuals. The use of SMC overcomes the drawbacks of the previously employed techniques

[2] in the number of nodes and the number of edges in the network


1
2
5
3
4
$\mathcal{E} = \{ 2 , 4 , 5 , 5 , 4 , 3 , \_ \}$
$Idx = \{ 1 , 4 , 5 , 6 , 6 , 7 \}$

(a) (b)

Figure 2: (a) A graph on five nodes and six edges; (b) The edgelist representation for the graph in Figure 2(a). It is important to note that $\mathcal{E}$ contains a dummy entry _ at the end to denote end of list. Idx list also contains a dummy node at the end for implementation reasons.

(like surveys, interviews and sampling) for studying distributedly held networks. In order to put the idea to practice, we presented a list of approaches that one can take while designing efficient solutions. These included the design of small sized circuits that can be evaluated using the state of the art generic SMC constructions, the use of RAM based secure computation as well as the use of efficient data-oblivious graph representations. We discussed some of the results in this direction, highlighting the appropriate techniques that have been adopted. As future work, we propose a thorough investigation of all the network measures such as community detection, degree distribution, betweenness centrality, eigenvector centrality, etc., in the SMC setting that are yet to be addressed. The objective of the thesis is to not only provide the theoretical design of secure protocols for computing network measures, but also to provide a complete protocol suite that implements them. This way, we wish to surpass the privacy issues that are currently an impediment to analyzing several networks.

REFERENCES

[1] 2018. Add Health. (2018). http://www.cpc.unc.edu/projects/addhealth

[2] Abdelrahaman Aly, Edouard Cuvelier, Sophie Mawet, Olivier Pereira, and Mathieu Van Vyve. 2013. Securely solving simple combinatorial graph problems. In *International Conference on Financial Cryptography and Data Security*. Springer, 239–257.

[3] Abdelrahaman Aly and Mathieu Van Vyve. 2014. Securely solving classical network flow problems. In *International Conference on Information Security and Cryptology*. Springer, 205–221.

[4] Gilad Asharov and Yehuda Lindell. 2017. A full proof of the BGW protocol for perfectly secure multiparty computation. *Journal of Cryptology* (2017), 58–151.

[5] Mikhail Atallah, Marina Bykova, Jiangtao Li, Keith Frikken, and Mercan Topkara. 2004. Private collaborative forecasting and benchmarking. In *Proceedings of the 2004 ACM workshop on Privacy in the electronic society*. ACM, 103–114.

[6] Vladimir Batagelj and Matjaz Zaversnik. 2003. An O (m) algorithm for cores decomposition of networks. *arXiv preprint cs/0310049* (2003).

[7] Peter S Bearman, James Moody, and Katherine Stovel. 2004. Chains of affection: The structure of adolescent romantic and sexual networks 1. *American journal of sociology* 110, 1 (2004), 44–91.

[8] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. 1988. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*. ACM, 1–10.

[9] Marina Blanton, Aaron Steele, and Mehrdad Alisagari. 2013. Data-oblivious graph algorithms for secure computation and outsourcing. In *Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security*. ACM, 207–218.

[10] Peter Bogetoft, Ivan Damgård, Thomas Jakobsen, Kurt Nielsen, Jakob Pagter, and Tomas Toft. 2006. A practical implementation of secure auctions based on multiparty integer computation. In *International Conference on Financial Cryptography and Data Security*. Springer, 142–147.

[11] Justin Brickell and Vitaly Shmatikov. 2005. Privacy-preserving graph algorithms in the semi-honest model. In *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 236–252.

[12] David Chaum, Claude Crépeau, and Ivan Damgard. 1988. Multiparty unconditionally secure protocols. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*. ACM, 11–19.

[13] Lin Chen and Richi Nayak. 2011. Social network analysis of an online dating network. In *Proceedings of the 5th International Conference on Communities and Technologies*. ACM, 41–49.

[14] Thomas Y Choi and Yunsook Hong. 2002. Unveiling the structure of supply networks: case studies in Honda, Acura, and DaimlerChrysler. *Journal of Operations Management* 20, 5 (2002), 469–493.

[15] Rob Cross, Nitin Nohria, and Andrew Parker. 2002. Six myths about informal networks-and how to overcome them. *MIT Sloan Management Review* 43, 3 (2002), 67.

[16] Ivan Damgård, Matthias Fitzi, Eike Kiltz, Jesper Buus Nielsen, and Tomas Toft. 2006. Unconditionally secure constant-rounds multi-party computation for equality, comparison, bits and exponentiation. In *Theory of Cryptography Conference*. Springer, 285–304.

[17] Ivan Damgard and Jesper Buus Nielsen. 2003. Universally composable efficient multiparty computation from threshold homomorphic encryption. In *Annual International Cryptology Conference*. Springer, 247–264.

[18] Gilbert Fridgen and Tirazheh Zare Garizy. 2015. Supply Chain Network Risk Analysis-A Privacy Preserving Approach.. In *ECIS*.

[19] David F Gleich. 2015. PageRank beyond the Web. *SIAM Rev.* 57, 3 (2015), 321–363.

[20] Oded Goldreich, Silvio Micali, and Avi Wigderson. 1987. How to play any mental game. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*. ACM, 218–229.

[21] S Dov Gordon, Jonathan Katz, Vladimir Kolesnikov, Fernando Krell, Tal Malkin, Mariana Raykova, and Yevgeniy Vahlis. 2012. Secure two-party computation in sublinear (amortized) time. In *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 513–524.

[22] Florian Kerschbaum and Andreas Schaad. 2008. Privacy-preserving social network analysis for criminal investigations. In *Proceedings of the 7th ACM workshop on Privacy in the electronic society*. ACM, 9–14.

[23] Yusoon Kim, Thomas Y Choi, Tingting Yan, and Kevin Dooley. 2011. Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management* 29, 3 (2011), 194–211.

[24] David Krackhardt and Jeffrey R Hanson. 2001. 16 Informal Networks: The Company Behind the Chart. *Creative Management* (2001), 202.

[25] David Krackhardt and Robert N Stern. 1988. Informal networks and organizational crises: An experimental simulation. *Social psychology quarterly* (1988), 123–140.

[26] Varsha Bhat Kukkala and S. R. S. Iyengar. 2017. MPC meets SNA: A Privacy Preserving Analysis of Distributed Sensitive Social Networks. *CoRR* abs/1705.06929 (2017). http://arxiv.org/abs/1705.06929

[27] Varsha Bhat Kukkala, Jaspal Singh Saini, and SRS Iyengar. 2016. Privacy preserving network analysis of distributed social networks. In *Information Systems Security*. Springer, 336–355.

[28] Varsha Bhat Kukkala, Jaspal Singh Saini, and SRS Iyengar. 2017. Secure Multiparty Construction of a Distributed Social Network. In *Proceedings of the 18th International Conference on Distributed Computing and Networking*. ACM, 12.

[29] Yehuda Lindell and Benny Pinkas. 2009. A proof of security of Yao's protocol for two-party computation. *Journal of Cryptology* 22, 2 (2009), 161–188.

[30] Takashi Nishide and Kazuo Ohta. 2007. Multiparty computation for interval, equality, and comparison without bit-decomposition protocol. In *International Workshop on Public Key Cryptography*. Springer, 343–360.

[31] Luis EC Rocha, Fredrik Liljeros, and Petter Holme. 2010. Information dynamics shape the sexual networks of Internet-mediated prostitution. *Proceedings of the National Academy of Sciences* 107, 13 (2010), 5706–5711.

[32] Elaine Shi, T Chan, Emil Stefanov, and Mingfei Li. 2011. Oblivious RAM with O ((logN) 3) worst-case cost. *Advances in Cryptology–ASIACRYPT 2011* (2011), 197–214.

[33] Tamir Tassa and Francesco Bonchi. 2014. Privacy Preserving Estimation of Social Influence.. In *EDBT*. 559–570.

[34] Xiao Wang, Hubert Chan, and Elaine Shi. 2015. Circuit oram: On tightness of the goldreich-ostrovsky lower bound. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM, 850–861.

[35] Barbara K Wichmann, Barbara K Wichmann, Lutz Kaufmann, and Lutz Kaufmann. 2016. Social network analysis in supply chain management research. *International Journal of Physical Distribution & Logistics Management* 46, 8 (2016), 740–762.

[36] Andrew Chi-Chih Yao. 1986. How to generate and exchange secrets. In *Foundations of Computer Science, 1986., 27th Annual Symposium on*. IEEE, 162–167.

[37] Samee Zahur, Xiao Wang, Mariana Raykova, Adrià Gascón, Jack Doerner, David Evans, and Jonathan Katz. 2016. Revisiting square-root ORAM: efficient random access in multi-party computation. In *Security and Privacy (SP), 2016 IEEE Symposium on*. IEEE, 218–234.