International Journal For Technological Research In Engineering
Volume 2, Issue 11, July-2015 *ISSN (Online): 2347 - 4718*

HEURISTIC ANONYMIZATION SCHEME FOR SECURE SOCIAL NETWORKS IN CLOUD

V. Srikanth[1], Dr. Siddhartha Ghosh[2]
[1]M.Tech Student, [2]Professor and HOD
Department of CSE, Keshav Memorial Institute of Technology, Hyderabad, Telangana state, India.

Abstract: Cloud computing as an incipient commercial paradigm enables organizations that host gregarious network data to outsource a portion of their data to a cloud. Security and privacy are major issues in cloud computing. In the authentic world, companies would publish convivial networks to a third party, e.g., a cloud accommodation provider, for marketing reasons. We recognize the novel type of privacy attack, called as 1*-neighborhood attack. If suppose we surmise that an assailant has erudition about the degrees of a target's one-hop neighbors, in integration to the target's 1-neighborhood graph, it having the one-hop neighbors of the target and the relationships among these neighbors.

An attacker having information may have a chance to re – identify the target from a k-anonymity gregarious network it's probably higher than the 1/k, where any node's 1-neighborhood graph is isomorphic with k−1other nodes' graphs. So we introduce a key privacy property, probability indistinguishability, to the 1*-neighborhood attack in social network. We implement a heuristic indistinguishable group anonymization (HIGA) scheme to engender an anonymized convivial network with this privacy property. Finally it defines now also it's used give result to aggregate queries with high precision in social networks.
Index Terms: Indistinguishable, Cloud storage, Outsource, Anonymization, Social Network, Neighborhood.

I. INTRODUCTION

Cloud computing is a model of information processing, storage, and distribution in which highly centralized physical resources are furnished to remote clients on demand. Rather than purchasing authentic physical devices—servers, storage, and networking equipment—clients lease these resources from a cloud provider as an outsourced accommodation that abstracts away physical contrivances. By sharing infrastructure among holders, a cloud vendor got economies of scale and balances workloads, reducing per-unit resource costs and giving clients the ability to rachet their resource consumption up or down. Cloud computing is flexible and portable in that it can be accessed anytime from anywhere. By utilizing redundant sites and backup storage, cloud providers can additionally provide more preponderant reliability than local computing systems. For all the benefits of cloud computing, though, it deprives clients of direct control over the systems that manage their data.

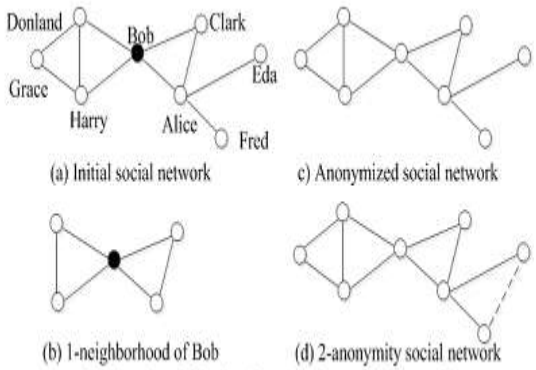


Fig. 1. 1-neighborhood attacks in a social network.

Convivial networks model convivial relationships with a graph structure utilizing nodes and edges, where nodes model individual convivial actors in a network, and edges model relationships between convivial actors. The relationships between gregarious actors are often private, and directly outsourcing the gregarious networks to a cloud may result in unacceptable disclosures. For example, publishing gregarious network data that represent a set of gregarious actors cognate by sexual contacts or shared drug injections may compromise the privacy of the gregarious Fig-neighborhood attacks in a gregarious network. Actors involved. Therefore, subsisting research has proposed to anonymize convivial networks afore outsourcing.

A native approach is to simply anonymize the identity of the convivial actors afore outsourcing. However, an assailer that has some cognizance about a target's neighborhood, specially a one-hop neighborhood, now also re-identifies the target with more confidence. This assailment, termed1-neighborhood attack, is proposed by Consider a synthetic gregarious network of "co-authors", as shown in Fig. 1-(a), where a node denotes an author, and an edge that links two authors denotes that they aforetime on a paper. In the neighborhood attack, an assailer, who kens Bob's one-hop neighbors and the connections between them, i.e., Bob's1-neighborhood graph, as shown in Fig. 1-(b), can still re-identify Bob from an anonymized graph, Fig. 1-(c), where all utilizer identities are abstracted. This is because Bob's 1-neighborhood graph is unique. To mitigate this assailment, Zhou et al. defined a k-anonymity gregarious network, where an assailant, with the erudition of any target's 1-neighborhood graph, cannot re-identify the target with confidence higher than1/k. Their rudimentary conception is to make any node's 1-neighborhood graph isomorphic with

at leastk−1other nodes' graphs by integrating noise edges. Given k isomorphic 1-neighborhood graphs, everyone has a probability of1/k to the target. For example, by integrating an edge between Eda and Fred, Fig. 1-(d) becomes a 2-anonymity convivial network.

In this paper, we identify a novel type of privacy attack, termed 1*-neighborhood attack, where an assailant is concluded to the ken degrees of the target's one-hop neighbors, in integration to the structure of the 1-neighborhood graph. We call this kind of background cognizance the1*-neighborhood graph. This postulation is plausible, since once the assailer kens the identities of the target's one-hop neighbors; he will be very liable to a mass more information about the one-hop neighbors, rather than only accumulating the connection information between them.

With this postulation, the assailant may re-identify the target from a k-anonymity gregarious network with a probability higher than1/k. To illustrate, let us postulate that the assailant kens the degrees of Bob's one-hop neighbors, Alice, Clark, Donland, and Harry, verbally express 4, 2, 3, 3, respectively. In Fig. 1-(d), the degrees of Alice's one-hop neighbors, Bob, Clark, Eda, and Fred, are 4, 2, 2, 2, respectively. Since only integrates edges to make 1-neighborhood graphs isomorphic, Alice can be omitted from the target candidate set, and the probability to re-identify Bob is 1. To deal with the 1*-neighborhood attack, requires the additament of many edges, so that the degrees of the k isomorphic graphs are identically tantamount. For example, by integrating edges between Grace and Fred, and between Grace and Eda, the degrees of Alice's one-hop neighbors are equipollent to that of Bob's. However, as more edges are integrated, the utilization of the gregarious networks will be further compromised. To sanction subsidiary analysis on the convivial networks, while preserving the privacy of the gregarious factors involved, we define a key privacy property, probabilistic indistinguishability, for an outsourced convivial network. To engender an anonymized gregarious network with such a property, we implement a heuristic indistinguishable group anonymization (HIGA) scheme. Our rudimentary conception consists of four key steps:

Grouping, we group nodes whose 1*-neighborhood graphs satiate certain metrics together, and provide an amalgamation and splitting mechanism to make each group size at least equipollent to k; Testing, in a group, we utilize desultory walk (RW) to test whether the 1-neighborhood graphs of any dyad of nodes approximately match or not; Anonymization, we implement a heuristic anonymization algorithm to do any node's 1-neighborhood graph approximately match those of other nodes in a group, by either combining or abstracting edges ;Randomization, we arbitrarily change the graph structure with a certain probability to ascertain each 1*-neighborhood graph has a certain probability of being different from the pristine one. Our contributions are threefold:

- We detect a novel attack, 1*-neighborhood attack, for outsourcing gregarious networks to a cloud.
- We develop the probabilistic indistinguishability property for an outsourced convivial network, and implement a heuristic indistinguishable group anonymization scheme (HIGA) to engender gregarious networks with this privacy property.
- We conduct experiments on both synthetic and authentic data sets to check the efficacy of the developed scheme.

II. CLOUD SERVICES

Cloud computing is anything that involves accommodations over the cyber world. These accommodations are broadly relegated into three categories: software as an accommodation (SaaS), platform as an accommodation (PaaS) and infrastructure as an accommodation (IaaS). Cloud software as an accommodation (SaaS) is the on-demand accommodation developed for end users; provider will license the software for their own use. As the software is managed over the central location over the web, the utilizer need not required to handle the upgrades. E.g: - Gmail. And the next accommodation is cloud platform as an accommodation(PaaS) is designed for the application developers, which provide all the facilities for developing the web applications facilely with more features without the involution of buying and maintaining the software and the infrastructure. E.g: -Google App Engine. Finally the cloud infrastructure as an accommodation (IaaS) is way of distributing the cloud computing infrastructure which provision the storage, accommodation and network. As it is planarity outsources accommodation it is not compulsory to buy the server, software and other equipments for the business and the accommodation providers advantage from cost preserving. E.g. - Amazon we accommodation.

III. RELATED WORK

In the following, we review the current state-of-the-art techniques and point out why they cannot plenarily address the "privacy-in-cloud" challenge. Privacy-Preserving Graph Publishing: Privacy auspice for graph publishing has been studied recently. Most of the subsisting work on graph publishing fixates on certain structural anonymizations, such as 1-neighborhood, k-degree, k-automorphism, k-isomorphism, and cluster predicated vertex anonymity, as well as many others. These techniques typically fixate on utilizing the least amount of modifications of the pristine graph (minimal information loss) to make it slake the targeted security requisite.

Lamentably, the anonymized graphs engendered from these privacy auspice techniques generally do not compulsorily maintain the statistical and graph theoretical characteristics of the pristine. In particular, for any dyad of vertices, there is no assurance of the degree of homogeneous attribute or preservation of shortest distances between the anonymized graph and the pristine graph. For example, k-isomorphism is proposed recently to partition a graph into k disjoint, isomorphic subgraphs, which lamentably cannot be acclimated to compute shortest distances in the pristine

graph. In integration, most of the exiting works deal with privacy on unweighted graphs, and do not consider the impact of edge weights. A few recent works notice the consequentiality of preserving graph theoretical characteristics during graph publishing. Ying and Wu propose to preserve the eigenvalue of a graph, which relates to average shortest distance and other topological features, during graph transformation. Liu et al. study edge weight perturbation by Guassian arbitrary or heuristic rules.

Das et al. propose a linear programming (LP) method to transmute edge weights while preserving shortest paths. The topological structure of the anonymized graph remains unchanged. Thus, even with the minimal topological erudition, such as the vertex degree, some sensitive information can be re-identified. Furthermore, the avaricious perturbation relies on an expensive matrix operation, and the LP approach can be facilely inundated by the number of inequality rules (shortest path preservation conditions). For example, for a connected graph with only one thousand vertices, there are one million rules for LP, which is pellucidly too expensive.

IV. PROPOSED SYSTEM

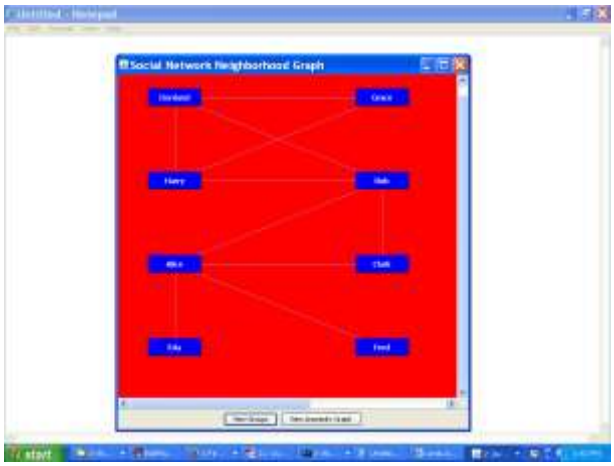
To sanction subsidiary analysis on the gregarious networks, while preserving the privacy of the convivial factors involved, we define a key privacy property, probabilistic indistinguishability, for an outsourced convivial network. To engender an anonymized convivial network with such a property, we introduce a heuristic indistinguishable group anonymization (HIGA) scheme. Our fundamental conception consists of four key steps: Grouping, we group nodes whose 1*-neighborhood graphs satiate certain advantages together, and provide a coalescence and splitting mechanism to make each group size at least equipollent to k; Testing, in a group, we utilize desultory walk (RW) to test whether the 1-neighborhood graphs of any dyad of nodes approximately match or not; Anonymization, we implement a heuristic anonymization algorithm to do any node's 1-neighborhood graph approximately match those of other nodes in a group, by either combining or abstracting edges ; Randomization, we arbitrarily change the graph structure with a certain probability to ascertain each 1*-neighborhood graph has a certain probability of being different from the pristine one.

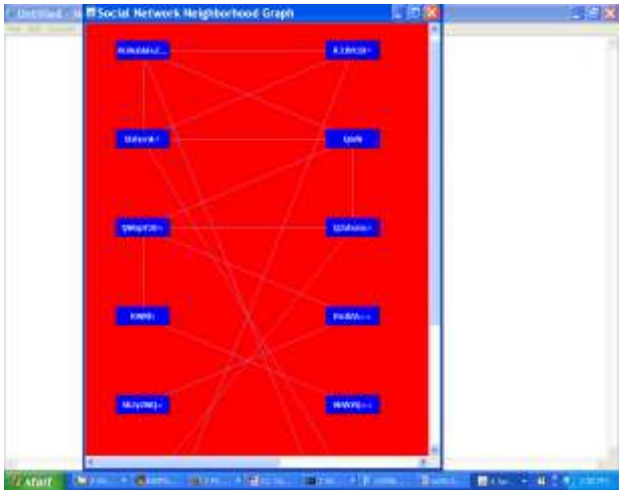
Advantages:

In this project, we identify a novel 1*-neighborhood attack. To resist this assailment, we introduce a key property, probabilistic indistinguishability for outsourced gregarious networks, and we propose a heuristic anonymization scheme to anonymize convivial networks with this property.

V. EXPERIMENTAL EVOLUTIONS

After admin logins, we can upload the dataset, here admin browse the file dataset.txt, file (data set) consists of list of users with their friends after clicking the submit button, the network neighborhood graph is drawn between the users as shown in the below screen, Here admin can view the Groups as well as Anonymity Graph.





After clicking the view group's button, groups will be formed. Here 4 groups are formed with user names Here Anonymity graph is drawn between the users i.e., more nodes and edges are added and also anonymization has done as shown in above screen.

VI. CONCLUSION

In cloud computing privacy preserving in Social Networks we define introduce a novel 1*-neighborhood attack. So we propose a key privacy property, probability in distinguish ability, to the 1*-neighborhood attack in social network. We implement a heuristic indistinguishable group anonymization (HIGA) scheme to engender an anonymized convivial network with this privacy property. Finally it defines now also it's used give result to aggregate queries with high precision in social networks.

VII. FUTURE ENHANCEMENT

We will expand our work in the future we will conduct an exhaustive theoretical study of jeopardies on the outsourcing convivial networks to a cloud, and endeavor to introduce other privacy mechanisms to our scheme, e.g., by cumulating with l-diversity, we enable the nodes in a group to be associated with at least l- different attributes. After that, the average node degree is 22 in the evaluation. However, in many convivial networks, the average node degree is much

higher which may make the proposed anonymization scheme inefficient. Therefore, on sizably voluminous convivial graphs with higher node density we will perform many experiments in the future.

REFERENCES

[1] D. Shasha, J.-L. Wang, and R. Giugno, "Algorithmics and applications of tree and graph searching," inProc. of PODS, 2002.
[2] X. Yan, P. S. Yu, and J. Han, "Graph indexing: a frequent structurebased approach," inProc. of SIGMOD, 2004.
[3] P. Zhao, J. X. Yu, and P. S. Yu, "Graph indexing: tree + delta >= graph," inProc. of VLDB, 2007.
[4] S. Zhang, M. Hu, and J. Yang, "Treepi: A novel graph indexing method," in Proc. of ICDE, 2007.
[5] J. Cheng, Y. Ke, W. Ng, and A. Lu, "Fg-index: towards verification-free query processing on graph databases," inProc. of SIGMOD, 2007.
[6] H. Shang, Y. Zhang, X. Lin, and J. X. Yu, "Taming verification hardness: an efficient algorithm for testing subgraph isomorphism," in Proc. Of VLDB, 2008.
[7] S. Berreti, A. Bimbo, and E. Vicario, "Efficient matching and indexing of graph models in content-based retrieval,"IEEE Trans. Pattern Analysis and Machine Intelligence, 2001
[8] Tran Khanh Dang. Privacy-preserving search and updates for outsourced tree-structured data on untrusted servers. In Proc. iTrust, 2005.
[9] Josep Domingo-Ferrer. A provably secure additive and multiplicative privacy homomorphism. In Proc. 5th International Conference on Information Security, 2002.
[10] W. Du and M. Atallah. Privacy-preserving cooperative statistical analysis. In Proceedings of the 17th Annual Computer Security Applications Conference, 2001.
[11] M. Kantarcioglu and C. Clifton. Privacy preserving k-nn classifier. In Proc. ICDE, 2005.
[12] A. Khoshgozaran and C. Shahabi. Blind evaluation of nearest neighbor queries using space transformation to preserve location privacy. In Proc. SSTD, 2007.
[13] Yaping Li and Minghua Chen. Privacy preserving joins. In Proc. ICDE, 2008.
[14] R. Chow et al. Controlling data in the cloud: Outsourcing computation without outsourcing control. In CCSW, 2009.
[15] S. Curry et al. Infrastructure security: Getting to the bottom of compliance in the cloud, March 2010. RSA Security Brief.

