ISSN 2393-8021

International Advanced Research Journal in Science, Engineering and Technology
Vol. 1, Issue 2, October 2014

From Facebook to Safebook : Review on Methods for Social Network Data Protection

S.Jeevitha[1], R.Santhya[2], Prof.S.Balamurugan[3], S.Charanyaa[4]

Department of IT, Kalaignar Karunanidhi Institute of Technology, Coimbatore, TamilNadu, India[1,2,3]

Senior Software Engineer Mainframe Technologies Former, Larsen and Tubro (L&T) Infotech, Chennai

TamilNadu, India[4]

Abstract: This paper reviews methods to protect social network data for the past 20 years. The social networking sites are the service providers who always used to be connected with peoples to share the information like photos, videos and personal messages. As the social networks usage grows, the risks behind them also increases. That makes the hackers, spammers, virus writers to attack the vulnerable nodes. Many users have problems for publishing posts and photos in a way that will protect them from the undesirable side effects to their online identities to support users' desire for "interactional privacy". Protection of the user's private sphere in online dealings with other people has suggested some improvements which have already been made for the networks such as Facebook. In a practical setting, however, these improved means are either too rigid to do justice to users' multifaceted habits or they are very complicated to manage because they try to solve a host of different problems all at the same time. An intruder may impose the privacy of the node with the help of published social network data and some background knowledge. For better privacy, the identities of the label such as social security number (SSN) of an employee, disease of a patient, etc., are replaced by some unique identity.

Keywords: Social Networking, Hackers, Virus, Facebook, Intruder, Privacy, Security

I. INTRODUCTION

The social network is useful in social sciences and researches to study the relationship between individuals or among societies. Inorder to share the data among individuals, there are various types of social networking websites to connect them. The social networks websites used at global level represent the place where people interact through online, make discussions, exchange photos or music or share their experiences. The fast development of this type of communication inturn gives rise to certain concerns towards the safety of using the Internet for the disclosure of some personal data. If it is decided to use this kind of social networking website and to post information about us, is said to be the creation of a profile of that particular person. Because of the information posted on the social networks websites become freely accessible to the public, an additional care to the information that disclosed about a particular person should be paid, which is simply because of the knowledge may be exposed to a large number of persons.

This knowledge disclosure can lead to privacy breach or even the person is physical or emotionally affected. A website which allows posting personal data must ensure the security of the information and also make sure that the data is not to be used for other purposes. This issue can be limited by making some changes with the websites. Most of these websites allow the setting of the information which is being posted as public or private information can be limited by using the control option "privacy settings". It is recommended to the particular person those using the social networks websites:

1) To avoid disclosing of personal information like the address or phone number.
2) To avoid using certain passwords which can be easily identified, even by closed persons like phone number, date of birth, closed one names etc.
3) To use a special e-mail address, different from the personal or professional one.

Cloud computing offers unique opportunities for supporting long-term record preservation. MyPHRMachines [1], a patient owned health record system prototype based on remote virtual machines hosted in the cloud. MyPHRMachines is particularly promising for countries with a very heterogeneous architecture of systems across hospitals and other care institutions. In the view of developer PHRs should be portable. PHR systems typically offer functionality to share, visualize and analyze PHR data. Secure lifelong management of patient medical records since data are stored in the cloud and do not have to be carried around by patients.

The remainder of the paper is organized as follows. Literature review of several techniques prevailing in protecting social network data security over past 20 years are discussed in Section 2. Section 3 concludes the paper and outlines the future work.

II. LITERARY REVIEW

Annemarie Mol and John Law [1994] [1] paper authors define what is social, what is network, about anemia and social topology. Some assumptions that frame the social performance were observed. Social performance can be similar or different. Authors describe anaemia with respect to three factors. First is 'region'. This outlines the things are bunched together and its limits should be drawn in every bunch. Second term is 'network'. It describes , connections between two origins and the space between origins. Final term is 'fluid spatiality' which describes alike or dislike at different fluid sites. Lastly the author says the remedies for fluid and anemia flow.

Richard H. Needle, Susan L. Coyle, Sander G. Genser and Robert T. Trotter II[1995] [2] authors demonstrated about drug use and HIV transmission on social network research. This paper deals about the reasons for people using drug and how HIV disease will transmit. It also says the remedies for problem. Also the author say the solutions for protect the drug users, by using network analysis and give the solution for HIV disease by changing the behavioral of persons who all are affected by disease.

Ravi Sandhu and Ierangela Samarti Author says[1996] [3] the three important factors for system to make a effective information system. They are Authentication, Access control and Audit. Authentication is related to user and system and it is kind of an interaction between user and system. Two techniques are used by author for authentication. First one is "Password", this is mostly used method. But it is extensively used by everyone and care should be taken while employing passwords. Another technique is "Token". Each token has different secret key. That key always come along with the token. If we try to separate or getting back the key, the key will be destroyed. "User-to-token" method is used here. In this method passwords are using in the form of Personal – Identification-Number(PIN).

The second term is access control ,which has control on what other party will do. This is the next process of authentication. Three types of access control is enlisted by author.

1) Discretionary access control(DAC)
2) Lattice based access control also known as mandatory access control(MAC)
3) Role based access control.

The final term is audit, which perform two tasks, collection and organization of audit data.

Laura Garton, caroline Haythornth waite Barry Wellman [1997] [4] in their paper spoke about how people are connected through social network. Author supports medium for communicating is Computer-Mediated-Communication (CMC).Because it focused on how individual users interact with computers ,as well as two persons and small groups. This paper also says the back-forth of social network approaches like CMC, computer network and social process. Social network is examined by relations among people, governments, according to situations. If people are gathering improper data, it will cause problems in data management, clarification and privacy.CMC are having many features like interacting with two persons in the form of software, hardware and 'groupware'. In case any different information is given by user by using CMC it is easy to find.

Adrienne Felt &David Evans[1998] [5] says about one of the social network "Facebook"and its privacy terms. The facebook user's data is accessed by third party developers. That third party developers may know the information like date-of-birth and gender. So, author introduced a privacy term called Privacy-by-Proxy. It will give a full privacy of user data. Due to this method the users does not concede their privacy.

Andreas Seufert, Georg von Krogh and Andrea Back [1999] [6] in their paper talked about knowledge management and networking. Both of two terms is very dominant for upcoming use in knowledge management. Authors also tells about the importance of networks in modern economy. Authors portray the relationship between network and knowledge management. They describe networks in knowledge management point of view. Authors say two types of method for knowledge management relationship. The first method is hypothetical basis of network. Second method is how network and knowledge management is depend on each another. This helps to introduce a new framework model called knowledge networking. This framework model expresses in micro perspective. This model describes as combining the group of people or resources. It contains three stages. First one is conditions, it should follow some facilitating conditions like management systems, corporate culture, organizational culture. The second level is working process of knowledge networking. It have resources like individual, group, organization, collectives of organizations etc., And last term is architecture of knowledge networking. It contains the tools for making structure, that tools are organizational tool and information and communication tool.

RakeshAgarwal and RamakrishnanSrikant[2000] [7] in their paper spoke about a productive track for data mining in the future work. Author construct a decision-tree, in that method contains the individual records. But it does not give exact number of records. Some other privacy preserving methods such as value-class membership, random distributions like Uniform, Gaussion were also discussed. The Uniform method is comfortable than Gaussian interms of operation. Under decision-tree author found two active algorithm namely Byclass and Local methods. Byclass method exhibits 25% -50% privacy. Local method exhibits higher privacy than byclass method.

Daniel J. Solove[2001] [8] said that privacy is created by the gathering and use of personal information through computer database and the internet metaphors with Big Brother. But author says this is not correct metaphor. In Big Brother method privacy law to protect the privacy is detailed. In this representation data privacy attack by shadowing, revelation of hidden information and discover the individuals hidden information were experimented. This method of privacy failed to give effective result. So author says kafka metaphor which solves the problem. It solve the problems like weakness, susceptibility.

Latanya Sweeney [2002] [9] talked about the necessity to share personal information in general place like hospital, bank. In that privacy preserving is very important. k-Anonymity is a technique coming for overcome this problem. It gives assurance to information privacy, when it is released. By using the following two terms for k-Anonymity method is described. One is Generalization, it is used for recoding the information. Second term is Suppression, it is used to prevent releasing the value. Authors used Minimal Generalization (MinGen) algorithm and above terms to make a minimal falsehood k-Anonymity for privacy protection. Author also says the about Datafly.

LeysiaPalen and Paul Dourish [2003] [10] says that in this networked world preserving privacy is a must. Every day new technologies evolve which increase the need for privacy. In particularly every new device invented and spread in market has new information and the privacy needs to be managed. Author construct a theory for unpacking privacy. It says privacy is vibrant and interaction process. Author discusses three strains for unpacking privacy and Human-Computer Interaction(HCI). Jennifer Golbeck and James.

Hendler Motive of this [2004] [11] paper is based on trust on semantic web. Authors introduced an algorithm for calculating fame rate. Also the authors give mathematical representation for the proposed algorithm.

Ralph Gross Alessandro Acquisti.H.John Heinz III [2005] [12] in their paper talked about information leak in social network and also its privacy involving in social network. The whole sum of information is expose in sizes. Also authors say about different view of privacy attacks. This paper especially deals about facebook and its various attacks such as physical and online.

Frederic Stutzman[2006] [13]author view the division of the same type of information in a network by using social network communities(SNC). In this paper author shows the result about identify information. Authors enumerates the methods used and the danger of identify information sharing in SNC. The most familiar is social network “Facebook” is taken into consideration. Author take a survey among colleges by users sharing personal information, political views which should be. Finally author says SNC is best compare with traditional method SSN(Social Security Number).

Katherine Straterand Heather Richter [2007] [14]in their paper dealt about one of the social network facebook and its privacy terms. Facebook have 18 million users, in which 80% to 90% of users are college undergraduates. Some other social networks dealt are Friendster and Myspace. Before constructing the privacy preserving technique privacy utilization and cooperation between users,users motivations and behaviors and trust of privacy should be concentrated. Privacy preserving techniques are increased to maintain a balance between utilization in social networks.

Craig E. Wills &Balachander Krishnamurthy[2008 (A)] [15] authors says about personal information protection on online social network(OSNs).This paper also talks about how information is exchanged and protection of information while sharing across OSN. OSN looks on possible privacy leakage. Authors compare OSN with other websites with respect to privacy leakage. Author also says about expansive of information sharing in this OSN. Author solution for privacy leakage is examined in the traditional website methods, and are implemented in OSN.

Bin Zhou, Jian PeiWo-Shun Luk [2008(B)] [16] paper authors talk about the use anonymization techniques for privacy protection. Theprivacy protecting technique is based on three parts background knowledge, privacy and data utility. Anonymization techniques are used for protecting data utility. Anonymization techniques used for protecting data,it is classified as two types.One is clustering based approaches and second method is graph modification approaches.

Matthew M. Lucas ,Nikita Borisov [2008 (C)] [17] authors say that social networks are very popular and used by many people. But social networks are having many privacy risks, If user send a message through social network, that message is detect and collect by service providers. To solve this problem the authors introduced a new architecture called flyByNight application. This favors facebook usability to protect the privacy without affecting the facebook usability. It houses certain features like encryption, to provide one-to-one communication and one-to-many communication.

Christina Prell and Klaus Hubacek and Mark Reed[2009(A)] [18] in their paper the author speak about the performance of the stakeholder analysis and social network analysis. Stake holders are used in natural resource management for taking decisions. For find out stakeholders , one must know the view of system. Top-down method is employed in spotting out the stakeholders.

LeucioAtonioCutillo ,RefixMolva,Thorsten Strufe, TU Darmstadt [2009(B)] [19] author talks about privacy and security in social networks. Author said social network should have the data honesty, availability and good data storage area. Author give an alternative for present social networks. That method called Safebook. Safebook exhibits data honesty and availability in distributed architecture.

Michael Zimmer [2010] [20] speaks about privacy protection in facebook. Authors mainly talks about future research of SNCs, factors for anonymization in public. To overcome of this problem author propose T3 method(Taste-Ties-Time).

Table 1 Analysis of Various Problems and Proposed Solution Methodology in Literature

S.No.	Year	Title	Problem Addressed	Proposed Solution Methodology
1.	1994	Regions, Networks &Fluids: Anemia& Social Topology	Anemia flow and Fluids.	Steps to safeguard from fluids.
2.	1995	Social networks, Drug Abuse, and HIV Transmission	HIV transmission.	changing the behavior of persons who all are affected by disease.
3.	1996	Authentication, Access control and Audit	To make effective system.	Authentication ,Access control and Audit
4.	1997	Studying Online Social networks	good medium for communicating	CMC is one of the good medium for communicating.
5.	1998	Privacy Protection for Social Networking APIs	Privacy protection in Facebook	Privacy-by-Proxy technique.
6.	1999	Towards knowledge Networking	Finding knowledge for future use.	A new framework model called knowledge networking.
7.	2000	Privacy-Preserving Data Mining	Privacy preserving in data mining.	Introduce decision tree method and also effective algorithms
8.	2001	Privacy and Power: Computer Databases and Metaphors for Information Privacy	Metaphors in information privacy.	Using kafka metaphor for information privacy.
9.	2002	Achieving k-Anonymity Privacy Protection Using Generalization and Suppression	Privacy preserving in government	Using MinGen algorithm.
10.	2003	Unpacking “Privacy” for a Networked World	Unpacking privacy in network.	HCI is one of the method for unpacking strains.
11.	2004	Accuracy of Metrics for Inferring Trust in Semantic Web-Based Social Networks	Trust on semantic web.	Solution for this problem is give algorithm and and mathematical representation.
12	2005	Information Revelation and Privacy in Online Social Networks	Information leak and privacy in social network.	Giving details of different methods for how privacy&leakage occur.
13.	2006	An Evaluation of Identity-Sharing Behavior in Social Network Communities	Identify the information in social network communities.	Taking survey among social network users.
14.	2007	Examining Privacy and Disclosure in a Social Networking Community	How privacy preserving terms are perform.	Giving the possibilities for making effective privacy preserving technique.
15.	2008(A)	Characterizing Privacy in Online Social Networks	Privacy Leakage	Study the traditional techniques expands and implement in OSNs.
16.	2008(B)	A Brief Survey on Anonymization Techniques for Privacy Preserving Publishing of Social Network Data	Privacy Protection	Anonymization techniques are used.
17.	2008(C)	flyByNight: Mitigating the Privacy Risks of Social Networking	Privacy Protection	flyByNight method used.
18.	2009(A)	Stackholder Analysis and Social Network Analysis in Natural Resource Management	Performance of stakeholder analysis and social networks.	Top-down strategy
19.	2009(B)	Safebook: A Privacy-Preserving Online Social Network Leveraging on Real-Life Trust	Analysis of Privacy and honesty of data in social networks.	**Safebook**
20.	2010	‘‘But the data is already public’’: on the ethics of research in Facebook	Privacy protection in facebook	T3 method.(Taste-Ties-Time).

III. CONCLUSION

This paper reviews methods to protect social network data for the past 20 years. The social networking sites are the service providers who always used to be connected with peoples to share the information like photos, videos and personal messages. As the social networks usage grows, the risks behind them also increases. That makes the hackers, spammers, virus writers to attack the vulnerable nodes. Many users have problems for publishing posts and photos in a way that will protect them from the undesirable side effects to their online identities to support users' desire for "interactional privacy".

Protection of the user's private sphere in online dealings with other people has suggested some improvements which have already been made for the networks such as Facebook. In a practical setting, however, these improved means are either too rigid to do justice to users' multifaceted habits or they are very complicated to manage because they try to solve a host of different problems all at the same time.

An intruder may impose the privacy of the node with the help of published social network data and some background knowledge. For better privacy, the identities of the label such as social security number (SSN) of an employee, disease of a patient, etc., are replaced by some unique identity. This survey would promote a lot of novel research directions in social network data protection mechanisms.

REFERENCES

[1] Annemarie Mol, John Law,"Regions, Networks and Fludies: Anaemiaana Social Topology ", Social Studies of Science vol.24 ,No 4,(Nov.,1994) 641-647.

[2] Richard H. Needle et.al., "Social Networks, Drug Abuse, and HIV Transmission" 1995

[3] Ravi Sandhu, Pierangela Samarti, "Authentication,Access control and Audit"1996

[4] Laura Garton, Caroline Haythornthwaite, Barry Wellman,"Studying Online Social Networks"1997

[5] Adrienne Felt, David Evans, "Privacy Protection for Social Networking APIs", 1998

[6] Andreas Seufert, Georg von Krogh and Andrea Back "Towards knowledge Networking", 1999

[7] RakeshAgarwal and RamakrishnanSrikant, "Privacy-Preserving Data Mining" IBM Almaden Research center,650 Harry Road, San Jose CA 95120,2000.

[8] Daniel J. Solove, "Privacy and Power: Computer Databases and Metaphors for Information Privacy" 2001

[9] Latanya Sweeney," ACHIEVING k-ANONYMITY PRIVACY PROTECTION USING GENERALIZATION AND SUPPRESSION" 2002

[10] LeysiaPalen, "Unpacking "Privacy" for a Networked World", 2003

[11] Jennifer Golbeck and James Hendler "Accuracy of Metrics for Inferring Trust and Reputationin Semantic Web-Based Social Networks" 2004

[12] Ralph Gross, "Information Revelation and Privacy in Online Social Networks (The Facebook case), 2005

[13] Frederic Stutzman, "An Evaluation of Identity-Sharing Behaviorin Social Network Communities" 2006

[14] Katherine Strater and Heather Richter, "Examining Privacy and Disclosure in a Social Networking Community", 2007

[15] Craig E. Wills , "Characterizing Privacy in Online Social Networks", 2008(A).

[16] Bin Zhou, Jian Pei, WoShunLuk "A Brief Survey on Anonymization Techniques for PrivacyPreserving Publishing of Social Network Data", 2008(B)

[17] Matthew M. Lucas, Nikita Borisov, "flyByNight: Mitigating the Privacy Risks of Social Networking", 2008(C)

[18] Christina Prell, Klaus Hubacek and Mark Reed, "Stackholder Analysis and Social Network Analysis in Natural Resource", 2009(A)

[19] LeucioAtonioCutillo and RefixMolva, "Safebook: A Privacy-PreservingOnline Social Network Leveraging on Real-Life Trust", 2009(B)

[20] Michael Zimmer ''But the data is already public'': on the ethics of research in Facebook 2010.

[21] S.Balamurugan, S.Charanyaa, "Principles of Social Network Data Security" LAP Verlag, Germany, ISBN: 978-3-659-61207-7, 2014.

BIOGRAPHIES

S.Jeevitha and R.Santhya are currently pursuing their B.Tech. degree in Information Technology at KalaignarKarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India. Their areas of research interests include Network Security, Cloud Computing and Database Security.

Prof.S.Balamurugan obtained his B.Tech degree in Information Technology from P.S.G. College of Technology, Coimbatore, Tamil Nadu, India and M.Tech degree in Information Technology from Anna University, Tamil Nadu, India respectively. He is currently working towards his PhD degree in Information Technology at P.S.G. College of Technology, Tamil Nadu, India. At present he holds to his credit 50 papers International Journals and IEEE/ Elsevier International Conferences. He is currently working as Assistant Professor in the Department of Information Technology, Kalaignar Karunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India affiliated to Anna University TamilNadu, India. He is State Rank holder in schooling. He was University First Rank holder M.Tech. Semester Examinations at Anna University, Tamilnadu, India. He served as a Joint Secretary of IT Association, Department of Information Technology, PSG College of Technology, Coimbatore, Tamilnadu, India. He is the recipient of gold medal and certificate of merit for best journal publication by his host institution consecutively for 3 years. Some of his professional activities include invited Session Chair Person for two Conferences. He has guided 12 B.Tech projects and 2 M.Tech. projects. He has won a best paper award in International Conference. His areas of research interest accumulate in the areas of Data Privacy, Database Security, Object Modeling Techniques, and Cloud Computing. He is a life member of ISTE,CSI. He has authored a chapter in an International Book "Information Processing" published by I.K. International Publishing House Pvt. Ltd, New Delhi, India, 978-81-906942-4-7. He is the author of book titled "Principles of Social Network Data Security", ISBN: 978-3-659-61207-7.

S.Charanyaa obtained her B.Tech degree in Information Technology and her M.Tech degree in Information Technology from Anna University Chennai, Tamil Nadu, India. She was gold medalist in her B.Tech. degree program. She has to her credit 12 publications in various International Journals and Conferences. Some of her outstanding achievements at school level include School First Rank holder in 10th and 12th grade. She was working as Software Engineer at Larsen & Turbo Infotech, Chennai for 3 years where she got promoted as Senior Software Engineer and worked for another 2 years. She worked at different verticals and worked at many places including Denmark, Amsderdam handling versatile clients. She is also the recipient of best team player award for the year 2012 by L&T. Her areas of research interest accumulate in the areas of Database Security, Privacy Preserving Database, Object Modeling Techniques, and Cloud Computing. She is the author of book titled "Principles of Social Network Data Security", ISBN: 978-3-659-61207-7.