SURVEY ON AUTOMATIC DETECTION OF SENSITIVE ATTRIBUTE IN PRIVACY PRESERVED HADOOP ENVIRONMENT USING DATA MINING TECHNIQUES

[1]Ms. Anitha Murthy R, [2]Ms. Dhina Suresh
[1]Research Scholar, [2]Assistant Professor
Department of Computer Science,
St Joseph College of Arts and Science for Women, Hosur, TamilNadu

Abstract : This study has been focused on the automatic detection of sensitive attributes using data mining technique in Hadoop environment. We establish Privacy Preserved Hadoop Environment (PPHE) that automatically detects sensitive attribute using data mining techniques. The proposed PPHE considers Twitter to enable user to post messages. The content of the posted tweets are wide ranging and contains private information such as email addresses, mobile numbers, physical addresses, and date of births. The major purpose of this study is in three folds. At First, we authenticate each twitter user using the combined algorithm RSA and Elgamal Algorithm. In the second, we classify the tweets into private and non-private attributes by using Type-2 Fuzzy Logic System. In the third, we pursue the data suppression technique to compress private tweets.

Index Terms **-** Data mining, Hadoop environment, Twitter, RSA, Elgamal Algorithm, Type -2 Fuzzy Logic System.

1. Introduction

Due to the advancement and its importance, new sub-domain of data mining has been invented namely Privacy Preserving of Data Mining (PPDM) over the few decades. Researchers have proposed many techniques to provide privacy for the sensitive data while keeping usability of the data that are Anonymization, Data Perturbation and Differential Privacy. In the past, all of those approaches have failed to extract sensitive information from the given data. In the current scenario, providing privacy of an individual's is an important and challenging issue [2],[8]. Online Social Networks (OSN) has become highly popular, where users are more and more lured to reveal their private information. To balance privacy and utility, many privacy preserving approaches have been proposed which does not well meet users personalized requirements [9-10].

Most social networks based data sources such as Twitter, Facebook etc., have unstructured data and no analytics or processing tools can work directly on this unstructured data. Machine learning algorithms are presented to analyze numerical data exactly to know the structure of numbers [8]. Association rule hiding is a very popular technique for preserving user's privacy against malicious activities from online social networks users. Sensitive rules and non-sensitive rules are formed which protect the sensitive knowledge from disclosure while sharing the data [2], [6].

Classification is a very important and necessary task in data mining. To identify the class label of unknown data, a more accurate classifier is required. Hence in [3] classification rule hiding is presented based on data distortion approach. However, to design a more accurate classifier, huge amount of data will be required. The private attributes of Twitter users can be inferred from tweets using text classification techniques. But the effective classification technique is necessitating for identification personal attributes [4].

Thus, the privacy preserving in data mining (PPDM) has emerged as a powerful component of data mining. The main aim of this work is to propose a framework that preserves the privacy of individuals.

2. Literature Survey

2.1 Privacy Issues in Social Networks and Analysis

Vu viet et al., [1] designate about the social networks since it generates numerous privacy issues for users in today's life. Hence, preserving privacy in social networks becomes a serious threat and there are several researchers have presented their researches related to this topic. This paper summaries the high level of social networks and determines their privacy threats and common privacy-preserving techniques which are used for solving privacy problems in social networks so that this paper categorizes the privacy issues in social networks into different groups in preserving users privacy.

2.2 Particle Swarm Intelligence and Impact Factor based Privacy Preserving Association Rule Mining for Balancing Data Utility and Knowledge Privacy

Kalyani et al., [2] have intended to protect private association rules (sensitive) and reducing the number of non-private association rules (non-sensitive) in database. Privacy preserving association rule mining is an area in which data owner can protect private association rules (sensitive knowledge) from disclosure while sharing the data. To safeguard sensitive association rules, individual data values of a database must be altered. Therefore, privacy concerns must not compromise data utility. A methodology that optimally chooses and modifies the transactions of the database is required to balance privacy and utility. Particle swarm optimization is a meta-heuristic technique used for optimization. Hence, an approach with particle swarm intelligence is established to select a set of database transactions for alterations to minimize the number of non-sensitive association rules that are lost and to maintain high utility of the sanitized database without compromising on privacy concerns. The projected method for hiding association rules was assessed based on some performance parameters including utility of the transformed database.

2.3 Privacy-Preserving Classification Rule Mining for Balancing Data Utility and Knowledge Privacy Using Adapted Binary Firefly Algorithm

G. Kalyani et.al [3] have intended to propose privacy preserving classification rule mining for balancing data utility and knowledge privacy. Classification is a problem in data mining which constructs a model to classify the data and then recognize the class label of unknown data based on the constructed model. Huge amount of data is necessary to form a more accurate classifier. Sharing of data is one of the solutions to have enormous amount of data. When sharing the data among business associates, some sensitive patterns which can be derived from the data need not be revealed to the others. This situation raises a motivating issue of retaining the shared data with high quality by hiding some sensitive patterns. This paper reports the problem of classification rule hiding by projecting a novel method based on data distortion approach. To select the best possible way of altering the instances and then selecting the optimal instances which reduces the loss of non-sensitive classification rules, a computational intelligence technique binary firefly algorithm is adapted with necessary changes. The transformed data set will be shared to the others which reveal only non-sensitive knowledge.

2.4 A Utility Maximization Framework for Privacy Preservation of User Generated Content

Yi Fang et.al [4] have proposed utility maximization framework for privacy preservation of user generated content. This paper demonstrated that the private traits of individuals can be inferred from user-generated content by using text classification techniques. Specifically, we study three private attributes on Twitter users: religion, political leaning, and marital status. The ground truth labels of the private behaviors can be readily gathered from the Twitter bio-field. Based on the tweets posted by the users and their corresponding bios, text classification yields a high accuracy of identification of these personal attributes, which poses a great privacy risk on user-generated content. We further propose a constrained utility maximization framework for preserving user privacy. The objective of this paper is to maximize the utility of data when modifying the user-generated content, while degrading the prediction performance of the adversary. The KL divergence is minimized between the prior

knowledge about the private attribute and the posterior probability after seeing the user-generated data. Based on this proposed framework, we investigate several specific data sanitization operations for privacy preservation: add, delete, or replace words in the tweets. We derive the exact transformation of the data under each operation.

2.5 Towards Privacy Preserving User Targeting, Journal of Communications and Information Networks

Jinghua Jiang et al., [5] explore user's behavior information for instance geographic locations, search histories, and purchase behaviors. However, gathering such behavior information of users raises severe privacy threats. This paper provided a brief review of such privacy preserving user targeting approaches as well as the security threats faced by user targeting.

2.6 Protecting Online Social Networks Profiles by Hiding Sensitive Data Attributes

Hatem Abdulkader et al., [6] offered about the online social networks protection for millions of Internet users. A utility based association rule hiding algorithm is proposed for preserving privacy of user profiles against attacks form OSN users. In the first step, reconstruct user's profile attributes by assigning privacy level to each attribute and the second step association rule hiding algorithm is constructed based on utility of privacy setting. Profile attributes reconstruction is not an easy task because assigning each privacy level to each attribute is a much complicated than the other approaches.

2.7 A Deep Learning Approach for Privacy Preservation in Assisted Living

Ismini et al., [7] have proposed deep learning approach for privacy preservation in assisted living. Deep learning is a promising approach in machine learning research with significant success in current years. So far the applications of Deep Learning are being used in various industries such as healthcare data and ambient assisted living (AAL) environments. This paper used Long Short Term Memory (LSTM) Encoder-Decoder that is a principal component of deep learning. Encoder was used to get an encoded version of user information whereas the decoder was utilized to decode this encoded information based on the privacy rules defined by the user.

2.8 Semantics based Sensitive Topics Diffusion Detection Framework Towards Privacy Aware Online Social Networks

Chinnaiah Valliyammai et.al [8] have proposed Semantics-based sensitive topic diffusion detection framework towards privacy aware online social networks. The advent of sharing sensitive information via Online Social Networks (OSN) has jeopardized the user to the extent that the privacy of millions of OSN users could well be compromised, with their data openly available in the public domain. Evidently, users lack in data privacy and the access control mechanisms available to avoid the risk of disclosure. Therefore a framework that automatically preserves the user privacy to detect sensitive topic and minimize the risk of sensitive information disclosure risk beyond the current privacy sceneries offered by OSN service providers is required. This explores a three-fold sanitization framework which precisely detects sensitive topics semantically using statistical topic model scheme which incorporates standard knowledge bases for tagging the sensitive topics discovered. The interaction documents from location-of-interest are subjected to SSAR–LDA using Gibbs Sampling to identify sensitive topic clusters with high location entropy. The proposed experimental result shows, (i) the sensitive topic clusters are identified with very high accuracy, (ii) despite the redaction approach, which eliminate the sensitive term, our proposed scheme enhance the privacy preserving policy by replacing the sensitive terms with suitable hierarchical generalization fetched from knowledge bases (iii) the probability of Kullback–Leibler (KL) divergence between sensitive and generalized sanitization terms on Twitter, with negligible information disclosure risk is acceptable, and (iv) the sanitization carried out for 10 sensitive topics, from 4500 user posts of 790 Twitter users, demonstrated high precision and recall, which can be correlated with advanced privacy settings for OSN users in the near future.

2.9 Privacy Preserving Big Data Mining: Association Rule Hiding using Fuzzy Logic Approach

Golnar Assadat Afzali et.al [9] have proposed association rule hiding using fuzzy logic approach. Association Rule Mining (ARM) which has been widely used to cause potential threat towards privacy of data. These techniques are suggested to avoid the risk of sensitive knowledge leakage. In the modern era, it is necessary to use some optimization techniques for large volume of data. This article deliberates data anonymization technique that is

highly suitable for big data mining. To speed up the mining process, parallelization techniques are used in the proposed model. Association rule hiding performs based on the fuzzy logic approach. Rules with confidence levels to each association rule are computed. Thus the rules with less confidence values are removed from the database. Generally speaking, association rule mining process has taken high execution time and leads to poor storage optimization (maintains more copy of the data).

2.10 Scoring User's Privacy Disclosure across Multiple Online Social Networks

Erfan Aghasian et.al [10] had proposed Scoring user's privacy disclosure across multiple online social networks. Users in online social networking sites unknowingly disclose their sensitive information that aggravates the social and financial risks. Hence, to prevent the information loss and privacy exposure, users need find ways to quantify their privacy level based on their online social network data. The previous studies focus on measuring the privacy risk and disclosure, consider only a single source of data, neglecting the fact that users in general can have multiple social network accounts disclosing different sensitive information. In this paper, explores about approach that can help social media users to measure their Privacy Disclosure Score(PDS) based on information shared across multiple social networking sites. In particular, we identify the main factors that have impact on users privacy, namely, sensitivity and visibility, to obtain the final disclosure score for each user. By applying the statistical and fuzzy systems, we can specify the potential information loss for a user by using obtained PDS.

2.11 Privacy Leakage through Innocent Content Sharing in Online Social Networks

Maria Han Veiga et.al [11] have proposed privacy leakage through innocent content sharing in online social networks. The increased popularity and ubiquitous availability of online social networks and globalized Internet access have affected the way in which people share content. The information which users willingly make known on these platforms can be used for various commitments, from building consumer models for advertising, to inferring personal, potentially invasive, information. In this work, we use Twitter, Instagram and foursquare dataset to convey the idea that the content shared by users, especially when aggregated across platforms, can potentially disclose more information than it was originally intended. We performed two case studies: First, we perform user anonymization by mimicking the scenario of ending the identity of a user making anonymous posts within a group of users. Empirical evaluation on a sample of real-world social network suggests that cross-platform aggregation introduces significant performance gains in user identification. The second task demonstrated that it is possible to infer physical location visits of a user on the basis of shared Twitter and Instagram content. The paper proposed informativeness scoring function that estimates the relevance and novelty of a shared piece of information with respect to an inference task. This measure is validated using an active learning framework which chooses the most informative content to teach at a given point of time. Based on a large-scale data sample, we show that by doing this, we can attain an improved inference performance. In some cases, this performance exceeds even the use of the user's full timeline.

2.12 Joining User Profiles Across Online Social Networks: from the Perspective of an Adversary

Qiang Ma et.al [12] has proposed joining user profiles across online social networks: from the Perspective of an adversary. Being the anchor points for building social relationships in the cyber-space, online social networks (OSNs) play an integral part of modern people's life. Since different OSNs are designed to address specific social needs, people take part in multiple OSNs to cover different facets of their life. While the fragmented pieces of information about a user in each OSN may be of limited use, serious privacy issues arise if a sophisticated adversary pieces information together from multiple OSNs. To this end, we undertake the role of such an adversary and demonstrate the possibility of "splicing" user profiles across multiple OSNs and present associated security risks to users. In doing so, we develop a scalable and systematic profile joining scheme, splicer, that focuses on various aspects of profile attributes by simultaneously performing exact, quasi perfect and partial matches between pairs of profiles.

2.13 OLAP textual aggregation approach using the Google similarity distance

Mustapha Bouakkaz et.al [13] has proposed OLAP textual aggregation approach using the Google similarity distance. Data warehousing and online analytical processing (OLAP) are essential elements to decision support. In the case of textual data, decision support requires new tools, mainly textual aggregation functions, for better and faster high level analysis and decision making. Such tools will provide textual measures to users who wish to analyze documents online. In this paper, we propose a new aggregation function for textual data in an OLAP context based on the K-means method. This approach will highlight aggregates semantically richer than those provided by classical OLAP operators. The distance used in K-means is replaced by the Google similarity distance which takes into account the semantic similarity of keywords for their aggregation. The performance of our approach is analyzed and compared to other methods such as Top keywords, TOPIC, TuBE and BienCube.

2.14 A Novel Scheme for Abatement of Privacy Concern by Controlling the Reachability in Online Social Network

Moumita Samant et.al [14] have proposed novel scheme for abatement of privacy concern by controlling the reachability in Online Social Network. Online social networks (OSNs) play a significant role to exchange the gamut of information from user oriented personal information to global marketing. Due to the pervasiveness of OSNs, the security and privacy control of information creates a threat. Nowadays, the OSN is a platform of socializing which allows its users to take part in information spreading among its friend circle. The user-popularity grows linearly or exponentially depending upon how the users are connected to their friends as well as visibility-settings (like friends, friend-of-friend, and public) of each. In many cases, the users instead of the intended target disclose information to a wider audience being unaware of the data sharing and information flow policies through the social networks. To address this issue, we propose a dynamic model for social network based on the edge property of friend of relation using semantic web tools.

2.15 Privacy Preserving Social Network Data Publication

Jemal Abawajy et.al [15] has proposed Privacy Preserving Social Network Data Publication. The summary of online social networks (OSN) has transformed the way people connect and interact with each other as well as share information. OSN have led to a tremendous explosion of network-centric data that could be harvested for better understanding of interesting phenomena such as sociological and behavioral aspects of individuals or groups. As a result, online social network service operators are compelled to publish the social network data for use by third party consumers such as researchers and advertisers. As social network data publication is vulnerable to a wide variety of re-identification and disclosure attacks, developing privacy preserving mechanisms is an active research area. This paper presents a comprehensive survey of the recent developments in social networks data publishing privacy risks, attacks and privacy-preserving techniques. This survey present, various types of privacy attack and information exploited by adversaries to perpetrate privacy attacks on anonymizes social network data. We present an in-depth survey of the state-of-the-art privacy preserving techniques for social network data publishing, metrics for quantifying the anonymity level provided and information loss as well as challenges and new research directions. This survey helps readers understand the threats, various privacy preserving mechanisms and their vulnerabilities to privacy breach attacks in social network data publishing as well as observe common themes and future directions.

2.16 The impact of different forms of cognitive scarcity on online privacy disclosure Computers in Human Behavior

Giuseppe A. Veltri [16] has proposed impact of different forms of cognitive scarcity on online privacy disclosure. The way in which people manage information disclosure contributes to one of the biggest challenges of the information age – online privacy. The current study sheds a light on the privacy paradox, a gap between attitudes and behaviour, by exploring the role of cognitive scarcity in privacy disclosure behaviour. Using a large sample of the UK online general population (N=969), we conducted a Randomized Controlled Trial experiment to test the effect of two forms of induced cognitive scarcity: ego depletion and working memory load, on information disclosure levels. The proposed work results indicate a significant effect of both forms of scarcity on information disclosure in the direction of increasing the latter, even in the context of a generalized high disclosure.

2.17 Control your Facebook: An analysis of online privacy literacy

Miriam Bartsch et.al [17] has proposed Control your Facebook: An analysis of online privacy literacy. For an effective and responsible communication on social network sites (SNSs) users must decide between withholding and disclosing personal information. For this so-called privacy regulation, users need to have the respective skills in other words; they need to have online privacy literacy. This study discussed about factors that potentially contribute to and result from online privacy literacy. In an online questionnaire with 630 Facebook users, we found that people who spend more time on Facebook and who have changed their privacy settings more frequently reported to have more online privacy literacy. People with more online privacy literacy, in turn, felt more secure on Facebook and implemented more social privacy settings. A mediation analysis showed that time spend on Facebook and experience with privacy regulation did not per se increase safety and privacy behavior directly, stressing the importance of online privacy literacy as a mediator to a safe and privacy-enhancing online behavior. We conclude that Internet experience leads to more online privacy literacy, which fosters a more cautious privacy behavior on SNSs.

2.18 Towards Exploiting Social Networks for Detecting Epidemic Out breaks

Sergio Di Martino et.al [18] has proposed exploiting social networks for detecting epidemic outbreaks. Social networks are becoming a valuable source of information for applications in many domains. In particular, many studies have highlighted the potential of social networks for early detection of epidemic outbreaks, due to their capability to transmit information faster than traditional channels, thus leading to quicker reactions of public health officials. Anyhow, the most of these studies have investigated only one or two diseases, and consequently to date there is no study in the literature trying to investigate if and how different kinds of outbreaks may lead to different temporal dynamics of the messages exchanged over social networks. Furthermore, in case of a wide variability, it is not clear if it would be possible to define a single generic solution able to detect multiple epidemic outbreaks, or if specifically tailored approaches should be implemented for each disease. To get an insight into these open points, we collected a massive dataset, containing more than one hundred million Twitter messages from different countries, looking for those relevant for an early out break detection of multiple diseases. The collected results highlight that there is a significant variability in the temporal patterns of Twitter messages among different diseases. In this paper, we report on the main findings of this analysis, and we propose a set of steps to exploit social networks for early epidemic outbreaks, including a proper document model for the outbreaks, a Graphical User Interface for the public health officials, and the identification of suitable sources of information useful as ground truth for the assessment of outbreak detection algorithms.

As a result of various works held on sensitive attribute detection in privacy preserved Hadoop environment. However, detection of sensitive information in Hadoop environment still critical concern which has to be overwhelmed.

Table 2.1 Literature survey summary

Author/Date	Topic/Focus/Question	Concept/Theoretical method	Paradigm/ method	Findings	Limitations/ Gaps/Future Research
Vu Viet Hoang Pham., Shui Yu., Keshav Sood., Lei Cui., (2017)	Privacy Issues in Social Networks and Analysis: A Comprehensive Survey	Quadratic programming Tape framework Cryptography	Quantitative method	Common privacy metrics Privacy preserving techniques	To have a unified privacy measurement Privacy in the mobile social networks To increase the flexibility for multimedia co-owners to control postings.
G. Kalyani., M.V.P. Chandra SekharaRao., B. Janakiramaiah., (2017)	Privacy-Preserving Classification Rule Mining for Balancing Data Utility and Knowledge Privacy Using Adapted Binary Firefly Algorithm	Binary firefly algorithm Computational intelligence techniques	Quantitative method	The privacy of sensitive rules by maximizing the utility of the transformed data set when compared to the existing algorithms	To consider data set with more number of values and in numerical format
Yi Fang., Archana Godavarthy., Haibing Lu., (2016)	A Utility Maximization Framework for Privacy Preservation of User Generated Content,	Text classification techniques	Quantitative method	Private attributes of Twitter users can be accurately inferred by simple text categorization with the labels automatically extracted from the bio fields Specific data sanitization transformations are derived framework to preserve user privacy while maximizing data utility	Explore semantic preserving operations to produce meaningful tweets in practice Utility maximization framework with probabilistic inference To alarm users about the privacy risk and how to suggest the changes on the user

					generated content
Ismini Psychoula., Erinc Merdivan., Deepika Singh., Liming Chen., Feng Chen., (2018)	A Deep Learning Approach for Privacy Preservation in Assisted Living,	Long Short Term Memory Encoder-Decoder model	Quantitative method	To get an encoded version of user information To decode this encoded information according to privacy rules defined by the user.	The tokenization of the attributes to improve the performance The use of simulated data Expansion of the model to real world data, as well as more complex data formats such as meta-data and multimedia formats.
Hatem Abdulkader., EmadElAbd., WaleedEad., (2016)	Protecting Online Social Networks Profiles by Hiding Sensitive Data Attributes	Association rule hiding algorithm	Quantitative method	Proposed a framework for hiding sensitive data attributes in OSN user's profiles	Add the user recommendation preferences as a privacy utility to protect the user's profiles data.
Chinnaiah Valliyammai., Anbalagan Bhuvaneswari., (2017)	Semantics based Sensitive Topics Diffusion Detection Framework Towards Privacy Aware Online Social Networks	Statistical topic model scheme Gibbs Sampling	Quantitative method	Identify user posts with a sensitivity that exceeds the SIC threshold for various access level users Sanitizes user posts with sensible entities.	Extend our study for multi-linguistic user profiles on a wide variety of sensitive topics Posts that include images as well may be further morphed to accompany the sanitized description of the post

					cumulative framework can be regulated so that the cyber investigators address strategic decisions earlier on access points by minimizing the running cost in the course of sanitization
GolnarAssadatAfzali., ShahrirarMohammadi., (2017)	Privacy Preserving Big Data Mining: Association Rule Hiding using Fuzzy Logic Approach	Data anonymization technique Parallelization techniques Big data association rule hiding technique	Quantitative method	Hide sensitive rules. Undesired side effect of deleting frequent item-sets (ISs) on new entrance data is removed	To decrease undesired side effect of the proposed model to gain less information loss.
ErfanAghasian., SaurabhGarg, LongxiangGao., Shui Yu., James Montgomery., (2017)	Scoring User's Privacy Disclosure across Multiple Online Social Networks	Statistical and fuzzy systems	Quantitative method	Privacy measure of the users offers a positive perception for the users to have a more detailed examination of the information they want to share on multiple social networks	Generalisation of the privacy scoring framework considering users' perspectives about the sensitivity of their data explicit criterion for measuring the difficulty of data extraction
MoumitaSamanta1(B), Pinakpani Pal, and Abhik Mukherjee (2018)	A Novel Scheme for Abatement of Privacy Concern by Controlling the Reachability in Online Social Network	Semantic web tools	Quantitative method	Abatement of Privacy Concern by Controlling the Reachability in OSNs	The efficacy of the proposed control scheme should be tested with proper user feedback on other OSN as

					well.
Miriam Bartsch a, Tobias Dienlin, (2016)	Control your Facebook: An analysis of online privacy literacy	Psychometric study	Quantitative method	Factors that potentially contribute to and result from online privacy literacy.	Vertical online privacy literacy that is, how good are users at regulating access to their data for governments, providers, and institutions? Optimism bias, to use objective tests of online privacy literacy, such as the online privacy literacy scale

3. Conclusion

In recent years, popularity of Online Social Networks has rapidly grown. Protecting the privacy of users against unwanted disclosure poses challenging privacy problem. There are many research work proposed on privacy preserving in Hadoop environment using data mining techniques. In this paper, we explore the review of significant related works that are held on privacy preserving. The proposed techniques provide privacy for the sensitive data while keeping usability of the data that are Anonymization, Data Perturbation and Differential Privacy. But all those approaches have failed to extract sensitive information from the given data. The systematic review of existing privacy preservation techniques suggests a solution that would be best suitable for preserving the private information on user perception.

References

[1]. Vu Viet Hoang Pham., Shui Yu., Keshav Sood., Lei Cui., (2017). Privacy Issues in Social Networks and Analysis: A Comprehensive Survey, IET Networks, Special Issue on Security Architecture and Technologies for 5G, Vol. 7, Issue 2, PP. 74-84

[2]. G. Kalyani., M. V. P. Chandra Sekhara Rao., B. Janakiramaiah., (2017). Particle Swarm Intelligence and Impact Factor based Privacy Preserving Association Rule Mining for Balancing Data Utility and Knowledge Privacy, Arab J SciEng, PP. 1-18

[3]. G. Kalyani., M.V.P. Chandra Sekhara Rao., B. Janakiramaiah., (2017). Privacy-Preserving Classification Rule Mining for Balancing Data Utility and Knowledge Privacy Using Adapted Binary Firefly Algorithm, Arab J SciEng, PP. 1-23

[4]. Yi Fang., Archana Godavarthy., Haibing Lu., (2016). A Utility Maximization Framework for Privacy Preservation of User Generated Content, ACM Transactions

[5]. Jinghua Jiang., Yifeng Zheng., Zhenkui Shi., Jing Yao., Cong Wang., Xiaolin Gui., (2016). Towards Privacy Preserving User Targeting, Journal of Communications and Information Networks, Vol. 1, Issue 4, PP. 22-32

[6]. Hatem Abdulkader., Emad ElAbd., Waleed Ead., (2016). Protecting Online Social Networks Profiles by Hiding Sensitive Data Attributes, Procedia Computer Science, Vol. 82, PP. 20-27

[7]. Ismini Psychoula., Erinc Merdivan., Deepika Singh., Liming Chen., Feng Chen., (2018). A Deep Learning Approach for Privacy Preservation in Assisted Living, arXiv:1802.09359v1 [eess.SP] 22 Feb 2018
[8]. Chinnaiah Valliyammai., Anbalagan Bhuvaneswari., (2017). Semantics based Sensitive Topics Diffusion Detection Framework Towards Privacy Aware Online Social Networks, Cluster Computing, PP. 1-16
[9]. Golnar Assadat Afzali., Shahrirar Mohammadi., (2017). Privacy Preserving Big Data Mining: Association Rule Hiding using Fuzzy Logic Approach, IET Information Security, IET Journals, The Institution of Engineering and Technology, PP. 1-10
[10]. Erfan Aghasian., Saurabh Garg, Longxiang Gao., Shui Yu., James Montgomery., (2017). Scoring User's Privacy Disclosure across Multiple Online Social Networks, PP. 1-13
[11]. Maria Han Veiga, "Privacy Leakage through Innocent Content Sharing in Online Social Networks", ACM, 2016
[12]. Qiang Ma, "Joining User Profiles Across Online Social Networks: from the Perspective of an Adversary ", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2016
[13]. Mustapha Bouakkaz, "OLAP textual aggregation approach using the Google similarity distance", Int. J. Business Intelligence and Data Mining, PP-31-59, Vol. 11, No. 1, 2016
[14]. Moumita Samanta1(B), Pinakpani Pal, and Abhik Mukherjee , "A Novel Scheme for Abatement of Privacy Concern by Controlling the Reachability in Online Social Network" , Springer, 2018
[15]. Jemal Abawajy, Senior Member, IEEE, Mohd Izuan Hafez Ninggal and Tutut Herawan, "Privacy Preserving Social Network Data Publication", IEEE, 2016
[16]. Giuseppe A. Veltri, Andriy Ivchenko, "The impact of different forms of cognitive scarcity on online privacy disclosure", Computers in Human Behavior, 2017
[17]. Miriam Bartsch a, Tobias Dienlin, "Control your Facebook: An analysis of online privacy literacy", Elsevier, PP-147-154, 2016
[18]. Sergio Di Martino, Sara Romano, Michela Bertolotto, Nattiya Kanhabu, Antonino Mazzeo, Wolfgang Nejd, "Towards Exploiting Social Networks for Detecting Epidemic Out breaks", Springer, 2017