International Journal of Advanced Educational Research
ISSN: 2455-6157
Impact Factor: RJIF 5.12
www.educationjournal.org
Volume 3; Issue 1; January 2018; Page No. 56-58

Big data for address security and privacy issues in social networks

[1] Dr. Savita Kumari Sheoran, [2] Romika Yadav
[1] Department of Computer Science & Engineering, Indira Gandhi University Meerpur, Rewari, Haryana, India
[2] Research Scholar, Department of Computer Science & Engineering, Indira Gandhi University Meerpur, Rewari, Haryana, India

Abstract
The recent advent in the internet and web technologies has drastically change the pattern of our social interactions. Now a day, the more and more people are joining virtual societies instead of physical societies. This gargantuan use of social media store and share a lot of person information and track the digital footprint of the user. If this information is not handled properly by the social media platform owner and use, it may be stolen by anti social elements and hackers for their personal gain which may include fraud, spam, phishing, malware and third party application. Therefore, it is necessary that the information available on social media be handled and cares properly to avoid misuse. Apart from it, the social media platform generate enormous amount of data which are generally in the differentiated forms like structured, semi-structured and unstructured formats, which could not be handled properly with the traditional database management techniques. Big Data is the term recently coined to handle such data effectively. This paper intends to study the security and privacy issues in published text, media and web pages on the social media regime so that user can enjoy the benefits of these sites without fear of data threats and risks.

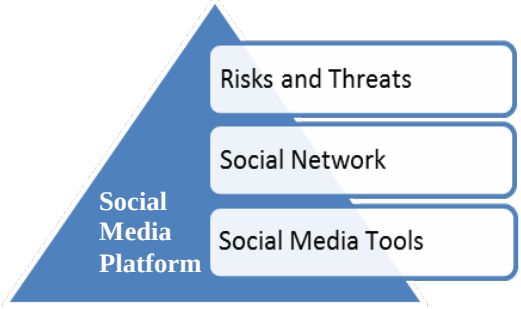
Keywords: social media, security and privacy, social networks and security threats

1. Introduction
Today's internet has approximately 3000 million users with 253 million available networking web pages. Due to accessibility of large data set various search engine will help to retrieve the valuable information from data of pools which is in millions. Google, Bing, Yahoo Search, Ask, My Web Search, Infospace, Aol Search, Contenko, Dogpile, Wow, WebCrawler, Info, DuckGo, Blekko, Alhea etc. are available search engines to get the correct information. Accordingly to success for WWW various technologies are applied with them to get the benefits of the internet. Users are connected with the social networking site which leads to connectivity of users. The popular social networking sites *viz.* Facebook, Twitter, Linkedln, Pinterest, Tumbler, Instagram, Google Plus, VK, Flickr, Tagged, Ask.fm, Meet Me, Class Mates, Vine and Meetup. The main aim of security in social networking sites to enjoy the benefits of these sites without fear from the data threats and risks. These risks and threats will affect the individual user to the whole organization. Large data set are exchanged by the users on the social networking sites. To handle these data sets become secure is a challenging task. In this paper we are introducing security strategies with the Big Data technologies, which is crux from the past in arrears to progression of technology and its potential to be applicable to variety of fields. Big Data sets help to analyze the various patterns for security and privacy issues. The data basically analyzed from the big data set to extract the knowledge of the big data analysts. Big Data analytics will have intense possessions across a range of fields *viz.* medical, health care, data mining, data warehousing, manufacturing, transportation, predictive analysis energy, utilities and social media. The rest of this paper is systematized as follows: section 2 survey the available literature in the area of security and privacy issues of social media. Section 3 constitutes the problem pertaining to the issues of social networking sites. Section 4 gives Social Network Security Risk, Malicious Attacks and Common Threats. Section 5 consists of analysis of security and privacy issues of social networks. Section 6 discusses the security tools and 7 concludes the paper.

2. Related Literature
Mohd and Jemal address the privacy of data published on the social networking sites. This paper defines data mining methods which are easier and efficient to handle the sensitive information. It also discuss privacy attacks on the social networks and provide state-an-art advances to preserve sensitive information using algorithms of sanitized data. Ahren *et al.* analyze the privacy decisions of photo sharing process of mobile users. It identifies the location of photo where it is captured corresponding to privacy settings of profile and endorsed the support users to privacy preferences for enhancing the awareness of information. Fang and LeFevere, only focused on privacy settings that help the user to find the appropriate privacy settings on the social networking sites. Mannan *et al.* addressed the problem on sharing data which people usually shares using their own web pages. They focused on privacy enabled services and instant messaging to create strategies. Besmer *et al.* discourses the photo tagged by person which allows the users to tag the request of owner to hide the picture link from the certain people. Squicciarini *et al.* accounts for the shared data that does not always belong to single user and propose the media item for establishing a collaborative management and shared content.

3. Motivation

An envisioned from the above section, a lot of work has been concluded by various researchers on the security and privacy of data on the social networking sites. Large networks are connected to each other through networked media. These privacy issues over the social media networks need to be analyzed to obtain amicable solution to overcome these issues of social media. Social network publish data periodically which is dynamic, evolutionary and huge in volume, ultimately falling to range of Big Data. The collective issues and strategies in the regime of social media are presented in figure below. Keeping all these factors at center this paper tried to calibrate inclusive plan to mitigate them.



4. Social network security risk and common threats

Risks are associated with the organizations those use the social networking sites for the personal reasons or authorized reasons. They are fatal to the organization if there is a major attack on their information; the association may leads to failure of organization. Also it is risks for the people who are associated with the social network to identify theft and even with their personal belongings. In 2009 a employee in Hawaii Hospital steal one of the patient information and posted the patient record with name and his medical confidential details on the MySpace Page. For this breach of confidentiality he was sentenced to one year jail for violating the law of health insurance portability and accountability. Social networking sites containing affluent information including email address, photos, date of birth, affiliations, and family information and personal addresses which is only required or needed for attackers. Attackers may use email address to send malicious data. Criminals or terrorist do not have any database regarding personal information of the user but they used the tools to automate search to extract confidential information on social networking sites. Due to popularity of social networking sites the attackers creates viruses and embedded viruses with spiteful links. When trying to sweeping these malicious links the computer gets affected and personal information of user may be deleted. Social engineering attacker may also collect personal information about work locations, education, photos, education and about family members. Attackers may access to credit cards and bank account information through e-mail or embezzle the naïve users through games and quizzes by providing malicious links.

5. Analysis of security and privacy issues

Social networking sites spawn so many applications and events and hence generate huge data which are venerable towards security risk. Basically the network traffic, system logs and other information are the some activities which are susceptible towards the threat. However, conventional security techniques having some issues in large data sets available to address these issues in social media but retain such a large dataset are not feasible for them. Most of the computer events, logs and their activities are deleted after a fixed period of time like 90 or 60 days. The available data is unstructured or semi structured and it is very difficult to analyze the complex queries with traditional techniques. Noise features also leads to inefficiency of datasets. Large datasets are expensive and their deployment needs relevant security attention. Privacy concerns are becoming more controversial and much publicized since the creations of social networking sites such as Facebook, Twitter, LinkedIn, Myspace and Bebo. The issues related to identify theft, checking personal information and stalking storage data using social media sites are more crucial to handle. A security issues arises when hackers access the unauthorized data from the protected websites. It will potentially harm the user's personal information, how much a user sharing its personal information on the social networking sites and how much user engaged in a social networking as well. Online social networks are now days very popular and claiming over million users. These users share their personal information, private content, photographs using social networks applications. Users must trust on the online social networks which provide benefits and examining the shared information. Several works that found the privacy issue found in the literature. For example author presented persona where users access their information on social network sites. Persona will hide the users data with the encryption attribute and apply fine grained policies for the users who wanted to view their data. Persona gives efficient and group based access mechanism which provide decryption and authentication to individual privacy used for the social networking platforms and protection of APIs that are used in online social networks. Many authors have proposed probabilistic neural-network privacy – preserving evaluating test algorithm in which data owned by other parties without knowing owner party. While Solanas proposed a NP-hard algorithm to solve privacy problem on data aggregation. Also suggested genetic algorithm for micro aggregation problems to solve them efficiently. These techniques provides opportunities with regard to strong and weak choices regarding to privacy protection on online social networking sites.

6. Security Tools

Social media faces a landscape of threats while communicating the data over the network. Advanced Persistent Threats (APTs) is erudite, multi-phase and extensive term which attacking to particular organization. Intrusion Detection System (IDS) and Security Information and Event Management (SIEM) that collects the information and aggregate them for presenting in actionable format for security analysis. New big data technologies such as Hadoop ecosystem, NoSQL databases, stream mining provides platform to analyse the security data on the large scale that also increase the speed and storage capacity beyond the traditional software. Big data having featured of multiple and unstructured data sets are used in single analysis framework. Big data technologies provide complete, less noisy, clean and

efficient data. Hadoop, Pig, Hive, Mahout, RHadoop, WINE Platform, Bot Cloud, MapReduce are used for complex events and efficiently processed the data for security purposes. Big data security model deliver unbelievable characteristics bases of data either internal or external to multiply the value and creating a synergistic effect, diverse data types collected through automated tools and normalize them accordingly.

7. Conclusion

With the span of time social networking sites are gaining more and more users and hence, it is on the radar of attackers. Many individuals and organizations are horribly affected with serious risks while using the social networking sites. This paper have analysed various security risks, malicious attacks and threats are discussed the strategies to mitigate them in ever changing digital world. Also with the heaping of record sets, the social media security and privacy issues increase. This paper have explored the applicability of Big Data technology with their characteristics to handle the privacy and security issues in social networks and found that such techniques posses enormous potential to address these issues effectively.

8. References

1. Abraham, Salama M *et al*., Computational Social Networks: Security and Privacy, springer - verlang London. 2012.
2. Carlos Otero E. Research Directions for Engineering Big Data Analytics Software, Published by the IEEE Computer Society. 2015, 13-17.
3. Alvaro Cardenas A. Big Data Analytics for Security, IEEE Computer and Reliability Societies. 2013, 74-76.
4. Mohd *et al*. Attack Vector Analysis and Privacy-Preserving Social Network Data Publishing, Proceedings conference IEEE computer society. 2011.
5. Moor JH. Towards a theory of privacy for the information age. SIGCAS Computer Society. 2010, 31-34.
6. Baden R *et al*. Persona: an online social network with user-defined privacy. ACM SIGCOMM Comput. Commun. Rev. 2009; 39(4):50-55.
7. Felt A, Evans D. Privacy Protection for Social Networking APIs. University of Virginia Charlottesville, Virginia. 2008.
8. Solanas A. Privacy protection with genetic algorithms. Stud. Computat. Intel. 2008; 92:215-237.
9. Barrigar J. Social network site privacy: a comparative analysis of six sites. The Office of the Privacy Commissioner of Canada. 2009.
10. Ahern S *et al*. Overexposed?: privacy patterns and considerations in online and mobile photo sharing. In Proceedings of the SIGCHI conference on Human factors in computing systems. 2007, 357-366.
11. Sood AK, Enbody RJ. Targeted Cyberattacks: A Superset of Advanced Persistent Threats, IEEE Security & Privacy. 2013; 11(1):54-61.