Risk Assessment in Online Social Networks Through Client Activity Analysis using Machine Learning

Sanaboina Chandra Sekhar

Assistant Professor, Department of Computer Science and Engineering, University College of Engineering Kakinada
Jawaharlal Nehru Technological University Kakinada
Email: *chandrasekhar.s[at]jntucek.ac.in*

Abstract: ***Online Social Networks (OSN's) are used to create a public or private profile that help in sharing information with other users and communicating with each other. These days OSN's are increasingly susceptible to privacy assaults that affect both users and their connections. In such cases risk assessment rating is given for each account. The idea is that malicious user behavior is different from normal behavior, which should be considered risky. As such, a particular basic behavioral pattern can't be described that suits all behaviors of OSN users. However, we expect related people to continue to obey common standards with related behavioral patterns. This research study introduces a twophase risk assessment technique, leveraging machine learning to identify atypical user behaviors that deviate from established norms. The initial phase groups similar users called as Group Identification Phase, and the second phase develops behavior models for these groups (Risk Assessment Phase). This study is carried out based on two features (i. e., Group Identifiers (GI) and Behavioral Features (BF)). The goal of this two phase risk assessment technique is to find most likely group for a given data set. The study applies KNearest Neighbors KNN for classification, categorizing users based on behavioral traits. Three groups were created and were named as Randomized Features (RF), Smart Risky User (SMR), and More Smart Risky Users (MSRU). The effectiveness of our risk assessment approach is measured based on three output metrics i. e., F - measure, detection rate and false alarm rate. This approach demonstrated notable improvements in risk assessment, with higher accuracy in detection rates and reduced false alarms.***

Keywords: Social Networks, Facebook, Twitter, Risk Assessment, Behavioral Analysis, Classification, K Nearest Neighbors

1. Introduction

Social networks remain in contact with families and friends for the purpose of exchanging personal knowledge and for business. Rapid increase in users in online media networks, privacy is at high risk. Users are not aware of privacy information. One of the key factors of these problems is that as a result of a huge volume of personal data being revealed, OSN users develop new connections with unfamiliar people. Users share the most confidential data on their accounts without defining suitable privacy controls and it can also lead to security issues. Such attacks can impact personal details of users, and also their friends' personal details.

Web - based social networks allow users to quickly create accounts and exchange information with other users. Such web sites of social networking bring people together that chat to each other, exchange thoughts and desires or make new friends. Given the popularity of OSN's, users have become a major issue for these social networking sites in preserving the data of the users and control over their data. When a person builds an account and starts to use a social media platform, they are in touch with one another, and certain persons they do not know, called strangers. OSN applications enable users to form new relationships with unknown individuals by exchanging large volumes of personal information. In these pages, the user interface shows a lot of confidential details that is available to everyone who wishes to see it. Sadly, Some users are less worried with privacy protection, and they share more confidential information on their accounts without defining the correct privacy settings and that can cause to privacy issues. Hence, exposure of personal information may lead to an action or occurrence of risk or an event for the personal information which can be devastating in some situations. A risk is an incident of low likelihood but it is difficult to analyze and has the high probability of major negative outcomes.

On the other hand, a event is an occurrence with greater likelihood, where the evidence is available to establish both the likelihood as well as the effects. Most people do not know how dangerous it is to share confidential details and the significant implications it might provide. Because some kind of attackers enjoys the benefits of social networking. Naive users trust in their fraud partnership/friendship. These fraudulent people use these naive users to spread spam and harmful material and to transmit messages over the system. Any of the details shared on such sites can also contribute to privacy threats, such as data fraud and cyberbully. In addition, certain types of hackers build fake identities and spread fake massages and harmful connections. In such situations, the message will be forwarded to his / her contacts once users press on such malicious items. Actual risky users send repetitive hurtful comments to their victims in some type of attacks like Cyber - bullying assault. All these kinds of hackers and fraudulent OSN account holders want to do these unsafe activities with all the other naive account holders and vendors posing a dangerous OSN situation.

1.1. Significance of this study

This study addresses the critical need for effective privacy measures in online social networks, proposing a two phase machine learning approach that enhances the detection of

Paper ID: SR241024183604 DOI: https://dx.doi.org/10.21275/SR241024183604

risky user behaviors, thus contributing to safer social networking environments.

1.2. Main aim of this article:

This article aims to develop and test a machine learning based risk assessment technique that identifies and categorizes risky behaviors within online social networks.

1.3. Risk Assessment

Risk assessment means assigning each user a risk rating by taking user actions and behaviors of closeness in the web. To detect abnormal habits, it's to associate the behavioral trends of users with other network users. One step risk assessment focuses only on personality characteristics. It is not possible to define all actions of OSN people using a common pattern, since related people seem to follow similar rules and behavior patterns. For this reason risk evaluation should be considering of two processing steps. This work proposes two step risk evaluation which focuses on group recognition based on both behavioral and personality characteristics i. e., related users are first grouped together depending on GI features, and we define one label for each defined community/group.

To lessen the overall impact of logistical threats, risk assessment is a continual process that involves defining, measuring, regulating, and evaluating learning. Even risk assessment is important with all the surroundings of their lives. Threat reduction helps communities minimize adverse effects and real risks and to remove or reduce risk exposures. Risk has become a part of any human now a days. Any of these hazardous behaviors may not be completely optional, because accidental risks are adverse impacts of an incident that occurs to us without our previous knowledge. Various activities are possible in online social networks, for example, writing comments and articles, reading or exchanging items, and specific forms of experiences, such as posting on new accounts, view the profile detail, allocate views, access particular groups or pages, submit invitations to anyone. The approach does not aim to monitor all users but focuses on identifying those who exhibit risky behaviors when creating a behavioral profile, but only those that expose risky behaviors. Write a message and post for instance getting any like on the post may lead to a situation where the subsequent user may become a target. On the other hand, it can not be considered a dangerous activity to simply own a large amount of contacts, tweets, reviews and shares. However it can be considered dangerous to have a large number of contacts, tweets, views and likes in a limited amount of time. The basic theory is that as the individual deviates against usual behavior the further it is considered dangerous behavior. Principle requires two main issues to be discussed. The first is the definition of a person's behavioral history capable of identifying certain user's habits and experiences which are required for risk evaluation. The second issue concerns the analysis of a 'normal behavior. ' In doing so, we have to assume that OSN applications of identified activities are genuinely heterogeneous. Because similar people tend to follow similar rules and have similar behavior. Depending on this concept, this research work suggests a two - phase risk evaluation as depicted in fig.1, where members were clustered/grouped/classified based on various attributes that are important for Group Identity (GI). Next phase is to create one or more typical behavior models for each group established.

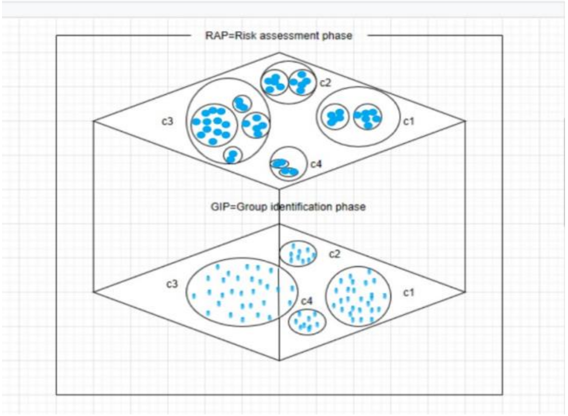

Figure 1: Two phase risk assessment

2. Review of Literature

Vitaliy Mezhuyev, S. M. Nazmus Sadat et. al in their paper **"Evaluation of the Likelihood of Friend Request Acceptance in Online Social Networks" [1]** discussed about Online social networks are example of a rash approval of a friend's invitation that can contribute to leakage of personal information and vulnerability to an attack. This paper suggests a way of estimating the probability of being a friend or a risk - free social network. Therefore all the incoming friend requests are evaluated with this model. This system helps user to filter requests from friends and cause alerts of odd behaviors.

Deveeshree Nayak, Summer Prince, et. al in their paper **"Information Privacy Risk Assessment of Facebook Graph Search" [2]** presents a risk assessment of Facebook graph. The number of users of social networking is increasing and the size of social networks is growing significantly. In this paper, we discussed data security risks associated with FB Graph Search. Build an FB Privacy Analysis platform where users can evaluate their own social networking security risk factors.

Michael Kaminsky et. al in their paper titled **"SybilLimit: A Near - Optimal Social Network Defense against Sybil Attacks" [3]** says that the Distributed, open networks (such as peer - to - peer services) are particularly susceptible to attack from Sybil. This paper introduces the novel procedure SybilLimit which utilizes a certain expertise as Sybil Guard but offers significantly enhanced and almost optimized commitments.

Naeimeh Laleh, Barbara Carminati, and Elena Ferrari et. al in their paper **"Graph Based Local Risk Estimation in Large Scale Online Social Networks" [4]** Online Social Networks (OSNs) are part of the daily lives of so many people. Our goal is to offer a risk score dependent on the variance of its heterogeneity variables to a particular person's close connections. This paper, First indicator of local risk assessment (Local risk factor) for directs interactions of the target customer. OSNs have some common characteristics in their social graph interface which vary from those of valid use.

Jose Alemanyi, Elena Del Val et. al in their paper "**Metrics for Privacy Assessment When Sharing Information in Online Social Networks" [5]**says that there are many users on social networking sites who are ignorant on security and also exchange information about who can use it. To suggest two metrics (Audience and Reachability) based on the limited information processes and relationship networks which indicate the risk of information sharing in terms of functionality. We proposed a new Security Threat model in this paper focused on relationship layers.

Raymond Heatherly, Murat Kantarcioglu, in their paper **"Preventing Private Information Inference Attacks on Social Networks" [6]** says that the data revealed within this network is meant to be confidential, but data can still be revealed in order to predict private information. The program has suggested different issues related to the personally identifiable information leakage. Here we use both friendship links and information together for greater accuracy, than using information alone. However, we also show that by removing data only, we greatly decrease the reliability of local classifiers, which gives us full detail.

Jinxue Zhang, Rui Zhang et. al in their paper **"The Rise of Social Botnets: Attacks and Counter measures" [7]** evaluates that OSN are increasingly thread by social boards. Social botnet refers to a group of network bots and one bot master power, merging one another to execute malicious activity. In this paper, digital influence manipulations on face book threw experiments and simulations to validate the merits of a botnet - based spam distribution, and also propose counter - measures to protect against these attacks and demonstrate their efficiency by simulating under real world data set.

Dan Yin, Yiran Shen, et. al in their paper **"Attribute Couplet Attacks and Privacy Preservation in Social Networks" [8]** discussed about the social network evaluation has dramatically changed the way we live. A number of attack model also been proposed recently. It uses the attribute couplets to assume identities within a social network. To propose k couplet annominity and develop the respective annonimization algorithms according to multiple public dataset evaluation.

Meghna Chaudharya and Harish Kumar B et. al in their paper **"Challenges in Protecting Personnel Information in Social Network Space" [9]** says that Social networking sites are now - a - days the lifeline for communication between people. The relations between nodes reflect interactors relationships. This paper discusses the complexities of shielding knowledge about employees in web environment. This paper discusses numerous cyber concerns including online harassment, web bashing and computer threats.

Jemal Abawajy, Tutut Herawan et. al in their paper **"Privacy Preserving Social Network Data Publication" [10]** describes OSN has resulted to a significant increase of network - centered data that could be collected to better understand interesting phenomena. This paper provides an in - depth analysis of the recent developments in dissemination of social networking data posing challenges to privacy, attacks and communication strategies. It analyzes and presents different types of attacks on privacy and information that attackers manipulate against publicly available social network data in order to conduct privacy attacks

3. Implementation

In social networks, risk assessment can be carried out in two stages: Group Identification phase and Risk Assessment phase.

3.1 Group Identification Phase:

First phase is group identification. Group identification features are age, gender and nationality, We must bear in mind that while individuals with a similar context normally behave in a way similar in the real world, in an OSN this can be influenced by the user's behavior against social networking, that might be specific only for related applications. The current work consists of different modules for the analysis of information on risk assessment. This phase involves implementation of Machine Learning models (both Supervised Machine Learning and Unsupervised Machine Learning). In any machine learning process most important phase is data preprocessing (Unsupervised Machine Learning) which involves preparation of the data for algorithm.

The K - Nearest Neighbors (kNN) algorithm is a popular and versatile machine learning algorithm for classification that has several advantages, including simplicity, adaptability, robustness, and versatility.

kNN can achieve high accuracy in a wide variety of prediction - type problems, especially when labeled data is too expensive or impossible to obtain.

3.1.1 Data Pre - processing

Pre - processing refers to transformations that are applied to our data before they are loaded into the algorithm. Many machine learning algorithms make assumptions with regard to your data. Machine learning algorithms often have a very good idea to arrange the data in such a way that the problem structure is better presented to them. Data Pre - processing is a way of converting original data into a clean dataset. In simple words, when the data is collected from a number of sources, it is collected in raw form which is not feasible for evaluation. Raw data is always incomplete (real - world data) and cannot be transmitted through a model. Data pre - processing involves various steps.

1) Import libraries
2) Load the data
3) Checking for categorical data
4) Checking for numeric data
5) Rescale the data
6) Splitting the data
7) Checking for missing and null values

3.2 Risk Assessment Phase:

The approach to risk assessment consists of two phases. First it seeks to coordinate members according to GI characteristics, the second manipulates Behavioral Features (BF).

Behind this purpose the attitude of users in online socialization is measured in addition to the personal information. Also features of GI include number of contacts, level of interaction and protection of profile. Second phase is based on behavioral features. Behavioral features are
Friendship rate (FR) =

$$FR\ (u)\ = \frac{|friends\ (u)\ |}{UserLongivity\ (u)}$$

Mutual friendship rate (MFR) =

$$MFR\ (u)\ = \frac{|mutualfriends\ (u)\ |}{friends\ (u)}$$

Friend mutual Friend rate (FMFR) =

$$FMFR\ (u)\ = \frac{|FR\ (u)\ |}{MFR\ (u)}$$

Comment Rate (CR) =

$$CR(u) = \frac{|CommentsBy(u)|}{UserLongivity(u)}$$

SC (Started Comments) =

$$SC\ (u)\ = \frac{|CommentsStartedBy\ (u)\ |}{UserLongivity\ (u)}$$

PR (Post Rate), PPS (Post Propagation Speed). These behavioral features are calculated by using two measures, The first is the lifespan of the client and is calculated as the amount of days the user has entered the OSN. The second one is the lifespan of the item, which is calculated as the amount of days after an item was entered into the OSN. Three models have created fake users to be inserted into the existing Facebook dataset. The first model identifies unsafe users with Randomized Features (RF), i. e. users who randomly set the GI attributes and BF's. The other includes Smart Risky users (SR), defined by relevant values for account characteristics (i. e. age, gender, employment, nationality), and conditional values for BF's. In the above model, More Intelligent Risky Users (MIRU) here false profiles are generated for valid identity and relationship attributes properties, and where the remaining behavioral characteristics are randomly selected. The attributes of communication are functions dependent on behavior and include: FR, MFR, and FMFR selected from BFs and number of friends selected from GIs.

To calculate the risk assessment over these two phases by using a algorithm as K - Nearest Neighbors (KNN). The KNN algorithm provides the similarities of features to determine the values of new information, which also means that the new data point should be given a value depending on how exactly it fits the training set points. We need data set for implementing any algorithm. During KNN's process we'll load the training as well as test details. First, they must select the K value, i. e. the closest data points to be considered. Any integer can be K. Using either form, measure the distance between the test data and each training data row namely: distance from Euclidean, Manhattan or Hamming formulas. Euclidean is the most commonly used tool for measuring distance. Next it will select the top K rows from the sorted range. Now, the test point is assigned a class label based on the most common class of those rows. KNN is straightforward to understand and interpret.

Risk assessment based on social network attacks create different test datasets for different attack forms (e. g., Sybils (dense graph), sybils (sparse graph), cyberbullying etc).

The dataset contains 5525 normal users and 1580 fake users for each dataset. Fake users are customized as per the attack being labeled. While the existing configurations are identical to regular users. Based on the behavioral patterns of sybil (dense friendship graph) the value of BFs such as CR (Comment Rate), SC (Started Comments), PR (Post Rate), PPS (Post Propagation Speed) can be anomalous, whereas other features will be similar to normal users. Experiments were performed on both the phases.

4. Results and Discussions

To perform the experiments on real data facebook dataset is used. The dataset was obtained using a Facebook framework that collected friendships and user knowledge. Dataset contains rows and columns of the following features like user id, user longevity, friends, mutual friends, and comments by, post by, likes by, comments started by, and profile secure, status. By using these features we can calculate the all behavioral features like friendship rate and mutual friendship rate etc. The data dimensions and data types of each feature must be known to handle a dataset very well. Initially, analysis of data is performed using the descriptive statistics and data visualizations to better understand the available dataset. Histograms and the matrix plot for correlation are the methods used for visualizing the results.

A histogram is an accurate, graphical representation of the numerical data distribution. Histograms are basically used to represent data given in the form of certain groups. X - axis is for bin ranges in which the Y - axis tells to frequency. The histograms for various parameters like CR, FR, FMFR, MFR were shown in fig 2.

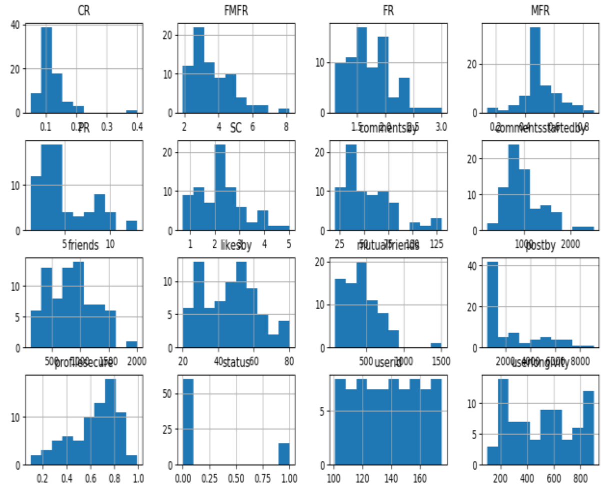

Figure 2: Histogram representation for statistical analysis of the dataset

The matrix of correlation is plotted to demonstrate which variable has a high or low correlation with respect to a different variable. Every variable in the diagonal line from top left to bottom right corresponds equally positively with each other. The correlation matrix for the implementation part is shown in fig 3 where 0, 1, 2, 3 ….13 represents various features that were considered in the dataset.

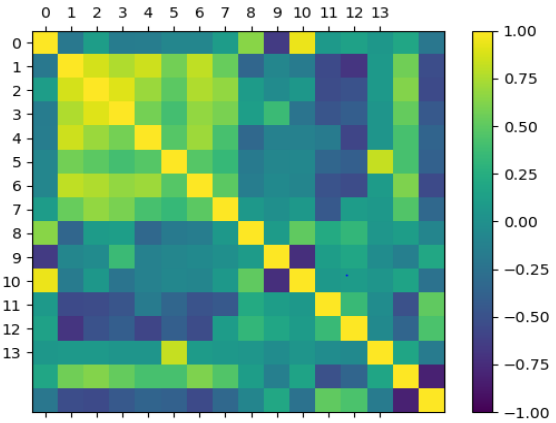

Figure 3: Correlation matrix for the facebook dataset

Prepare the data by removing duplicates, and test that there are no missing values in the dataset. The values of each feature in a data point will vary from random to normal. So, it's important to scale them in such a way that these rules are matches. The data is made up of attributes of different sizes, and many machine algorithms will benefit from rescaling the attributes to the same size for all. Features are often rescaled to the 0 - 1 level.

Accuracy of classification is the total number of correct predictions divided by the number of observations made for a dataset. The accuracy is inaccurate as a performance indicator for imbalanced classification issues. The accuracy of classification is widely used as it is a single measure used to describe results of the model. According to all three models our risk management approach is measured in terms of effectiveness.

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

F measure's calculated to our one phase and two phase risk assessment for three models RF (Randomized Features), MSR (Most Smart Risky Users), SR (Smart Risky Users) and their accuracy values as follows is depicted in fig 4.

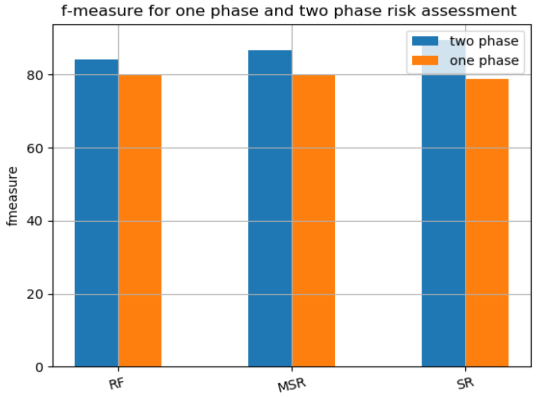

Figure 4: F - measure for one phase and two phase risk assessment

Detection rate is calculated to our one phase and two phase risk assessment for three models RF (Randomized Features), MSR (Most Smart Risky Users), SR (Smart Risky Users) and their accuracy values as follows is depicted in fig 5.

$$Detection\ Rate = True\ Detections / True\ Examples$$
$$= TP / (TP + FN)$$

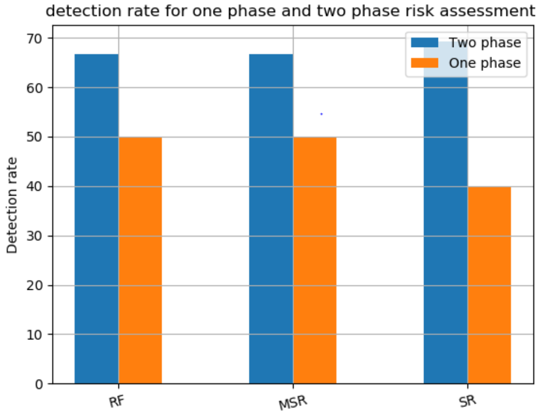

Figure 5: Detection Rate for one phase and two phase risk assessment

False Alarm Rate is calculated to our one phase and two phase risk assessment for three models RF (Randomized Features), MSR (Most Smart Risky Users), SR (Smart Risky Users) and their accuracy values are depicted in fig 6.

$$\boldsymbol{False\ Alarm\ Rate} = 1 - Specificity$$
$$= False\ Detections / All\ Detections$$
$$= FP / (TN + FP)$$

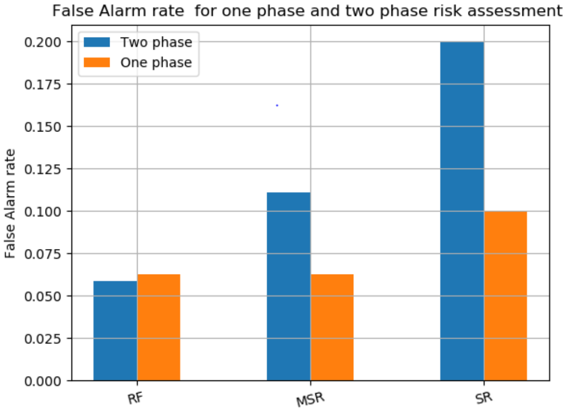

Figure 6: False alarm Rate for one phase and two phase risk assessment

5. Conclusion

This study proposes a two phase risk assessment framework using machine learning, focusing on behavior based user grouping and risk assessment in social networks. This method is capable of assigning each OSN user a risk score. This risk assessment is based on the premise that risky user behaviour is different from normal behaviour.

These risk estimations are done by using real Facebook dataset and the results showcases the effectiveness of our proposed work. The F - Scores, Detection Rates, and False Alarm Rates that were shown in the results section indicate that two phase risk assessment has achieved remarkable improvement in risk assessment compared to that of its counter part. In order to detect sybils, the study combined the features of supervised (KNN algorithm) and unsupervised methods (Scaling, Visualization, preprocessing etc.,) and evaluated the web usage behaviour of online users. Any of the other supervised behaviour - based risk models suffer from large false negative and positive levels due to the wide variety and unreliability of both normal and malicious OSN user's behaviour. Our findings indicate that this approach significantly enhances accuracy in identifying risky users and minimizes false alarms. Future studies could expand this framework for decentralized social networks, enhancing its adaptability and effectiveness.

6. Future Scope

Other machine learning models such as Decision Trees or Random Forest could also be considered for future work. In addition, this study can be extended to Decentralized Online Social Networks, which are distinguished by the lack of a specific data base to be evaluated.

References

[1] V. Mezhuyev, S. M. N. Sadat, M. A. Rahman, N. Refat, and A. T. Asyhari, "Evaluation of the Likelihood of Friend Request Acceptance in Online Social Networks, " *IEEE Access*, vol.7, pp.75318–75329, 2019, doi: 10.1109/ACCESS.2019.2921219.

[2] D. Nayak, S. Prince, and R. Robinson, "Information privacy risk assessment of Facebook Graph Search, " *Proc.2014 Sci. Inf. Conf. SAI 2014*, pp.1005–1006, 2014, doi: 10.1109/SAI.2014.6918309.

[3] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A near - optimal social network defense against sybil attacks, " *Proc. - IEEE Symp. Secur. Priv.*, no. Figure 1, pp.3–17, 2008, doi: 10.1109/SP.2008.13.

[4] N. Laleh, B. Carminati, and E. Ferrari, "Graph based local risk estimation in large scale online social networks, " *Proc. - 2015 IEEE Int. Conf. Smart City, SmartCity 2015, Held Jointly with 8th IEEE Int. Conf. Soc. Comput. Networking, Soc.2015, 5th IEEE Int. Conf. Sustain. Comput. Communic*, pp.528–535, 2015, doi: 10.1109/SmartCity.2015.124.

[5] J. Alemany, E. Del Val, J. M. Alberola, and A. Garcia - Fornes, "Metrics for Privacy Assessment When Sharing Information in Online Social Networks, " *IEEE Access*, vol.7, pp.143631–143645, 2019, doi: 10.1109/ACCESS.2019.2944723.

[6] R. Heatherly, M. Kantarcioglu, and B. Thuraisingham, "Preventing Private Information Inference Attacks on Social Networks, " *Tkde*, vol.25, no.8, pp.1849–1862, 2013.

[7] J. Zhang, R. Zhang, Y. Zhang, and G. Yan, "The rise of social botnets: Attacks and countermeasures, " *IEEE Trans. Dependable Secur. Comput.*, vol.15, no.6, pp.1068–1082, 2018, doi: 10.1109/TDSC.2016.2641441.

[8] D. Yin, Y. Shen, and C. Liu, "Attribute couplet attacks and privacy preservation in social networks, " *IEEE Access*, vol.5, pp.25295–25305, 2017, doi: 10.1109/ACCESS.2017.2769090.

[9] M. Chaudhary and H. Kumar, "Challenges in protecting personnel information in social network space, " *Proc.2015 Int. Conf. Emerg. Trends Networks Comput. Commun. ETNCC 2015*, pp.99–104, 2015, doi: 10.1109/ETNCC.2015.7184816.

[10] J. H. Abawajy, M. I. H. Ninggal, and T. Herawan, "Privacy preserving social network data publication, " *IEEE Commun. Surv. Tutorials*, vol.18, no.3, pp.1974–1997, 2016, doi: 10.1109/COMST.2016.2533668.