Identifying Vulnerable User In Linkedin Using Web Description Logic Rule Generation

Revathi.S, M.Suriakala

Abstract: One of the most preferred networks by professionals amid social networks is LinkedIn. The rapid and explosive growth of these social networks has enabled certain people to misuse the same for illegal and unethical conducts. Nonetheless, considering LinkedIn, these behavioral assertions prove very restrictive in the openly available profile information for users by privacy policies. The publicly present profile information of LinkedIn is limited. Here, it is suggested to pinpoint maximum group of the profile information required to identify vulnerable user in LinkedIn and also determine the proper data mining strategy for this task. In this paper Web Description Logic Rule Generation algorithm is put forth to find and examine vulnerable users and also used to remove the attackers from LinkedIn. Using this algorithm, identifying vulnerable users and to protect them against the attackers is possible according to the sharing threshold. When the threshold value exceeds the limit, the shared person will be removed from OSN. It is demonstrated that using limited profile information, this strategy is capable of spotting attackers at an accuracy of 94% and as low as 3.67% false negative.

Index Terms: Fake accounts detection, classification, LinkedIn accounts, Privacy content, monitoring, vulnerable, and malicious.

1 INTRODUCTION

Online social media networks like LinkedIn are witnessing huge increase in their user activity during the time that any event takes place in physical world. It is common that users upload personal photos and information on such websites without even being aware of about the sites' privacy. This paper recommends one system that has exclusively been developed for the application that focuses on detection of vulnerable user account in LinkedIn applications. In this paper new idea of proposed in Web Description Logic Rule Generation algorithm to find and examine vulnerable user and also attackers. Making use of this algorithm, it becomes possible to spot vulnerable users and attackers according to the sharing maximum information (or) posts and to remove the privacy attackers when found to exceed the threshold limit. It has been demonstrated that using sparse profile information, our method has the capability to pinpoint privacy attackers with an accuracy rate of 97% and only a small false negative of 2.67%. This is comparable to results gathered in other similar methods with big data set and also more profile data. In our proposed approach, we have succeeded in developing various stages of the social network building, uploading data, observing users' activity, machine learning approaches, identifying malicious user, and eliminating the vulnerable user. We have succeeded in establishing a set of exclusive features which help in distinguishing malicious applications from the real ones. This in turn would help reducing such malicious attacks. The goal of this project is observing the extent at which it will be possible by us to train our system for spotting properly the malicious users in LinkedIn applications on the OSN for achieving the final goal of reducing malicious privacy attacks. This paper has been organized in the below-mentioned manner: Section 2: Brief explanation of related works Section 3: Presenting the suggested machine learning strategies and features of various stages In section 4, results of tests have been shown. Lastly, in section 5, the paper concludes with proposal about research work in future.

- *Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author_name@mail.com*
- *Co-Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author_name@mail.com*

2 RELATED WORK

In the year 2014, Fire, M. et al. [1] attempted discovering spammers on the Twitter site by gathering a huge data set of Twitter OSN and categorizing the users into two: spammers and non-spammers. They used machine learning methods. Finally, they achieved success in pinpointing roughly 70% of the spammers while being able to identify 96% of the non-spammers. In 2013, Wald, R. et al. [2] took interest in pinpointing the malicious bots which was spreading harmful content or spam. In the end, they devised an automatic classification method. In the year 2013, S. Mahmood et al. [4] portrayed FIS that protects users of LinkedIn from attacks, This was in 2015, when the honeypot-oriented method was proposed by Gaurav Parsewar et al. [5] in order for discovering OSNs spammers. They succeeded in developing statistical user models using these honeypots for differentiating between legitimate users and social spammers. Similarly, in the year 2010, several "honey-profiles" were created by Stringhini et al. [6] on three huge OSNs, namely, MySpace, Twitter, and LinkedIn. They later examined the data that was collected and it was possible for them to pinpoint the abnormal behavior of users.

Adikari, Shalinda et al,[7]has succeeded spotting, the minimum set of profile information required to identify LinkedIn fake profiles and pinpoint the proper approach for data mining related to such processes. Only limited research with LinkedIn has been performed so far. [9]Hsieh et al. (2013), had performed a research about LinkedIn to understand the possibility of link among any two people in the organizational overlap. [10]Shuliang WANG et al. (2013) had made use of similarity and also interaction activity of the user profiles for developing a model which was not supervised for estimating the strength of friendship. M.R. Khayyambashi et al. (2013) have devised a system to discover the cloned profile using LinkedIn information. Result obtained from each query is considered by distiller for creating the user

record. Profile verifier which acts as the next component will examine the corresponding profile record to run similarity check against the original profile of the user. The profiles having the high possibility for cloning can finally be presented with the similarity ratings [11]. A multivariable model recognition method was used by Thomas et al. (2013) according to user profile name, email parameters, and screen name. Such model recognition method was able to be employed as pre-identifier to detect fake accounts before a detailed effort of authentication [12]. Several modern methods for spotting fake accounts in Twitter and also other OSNs focus on detecting grouped fake accounts based solely on their activity model. For instance, et al. (2014, 2016)[13] made use of an approach known as Catch Sync for detecting odd behavior in Twitter according to abnormal and synchronized user activity. In a similar way, [14] Clark et al. (2016) working a categorization method built on natural language coached on organic users for identifying messages sourced from automated accounts and for detecting fake accounts. (Xiao et al., 2015). Take for instance, El Azab et al. (2015) were capable of showing that the fake accounts in Twitter can be pinpointed with more efficiency that is based exclusively on a proven minimum set of parameters that includes the availability of the geo-information that used hash tags in the tweets, count of followers etc. [15] User's OSN action patterns [16] are reflected accurately in a social behavioral profile. Traditional desktops and also modern mobile devices are widely used by people for accessing OSNs.

3 PROPOSED WORK

3.1 overview

Social network websites like LinkedIn are seen to attract more than a huge 400 million users around the world. The growing popularity of these social networking leads to various issues that include the potential to Such a situation may lead to huge damage being caused in our real physical world, affecting the society in general, and particularly, business entities, citizens, and others. The paper has determined the minimum set of main factors which impact identification of vulnerability user in LinkedIn. It determines also factors that are applied by using Web Description Logic Rule Generation algorithm categorization methods.

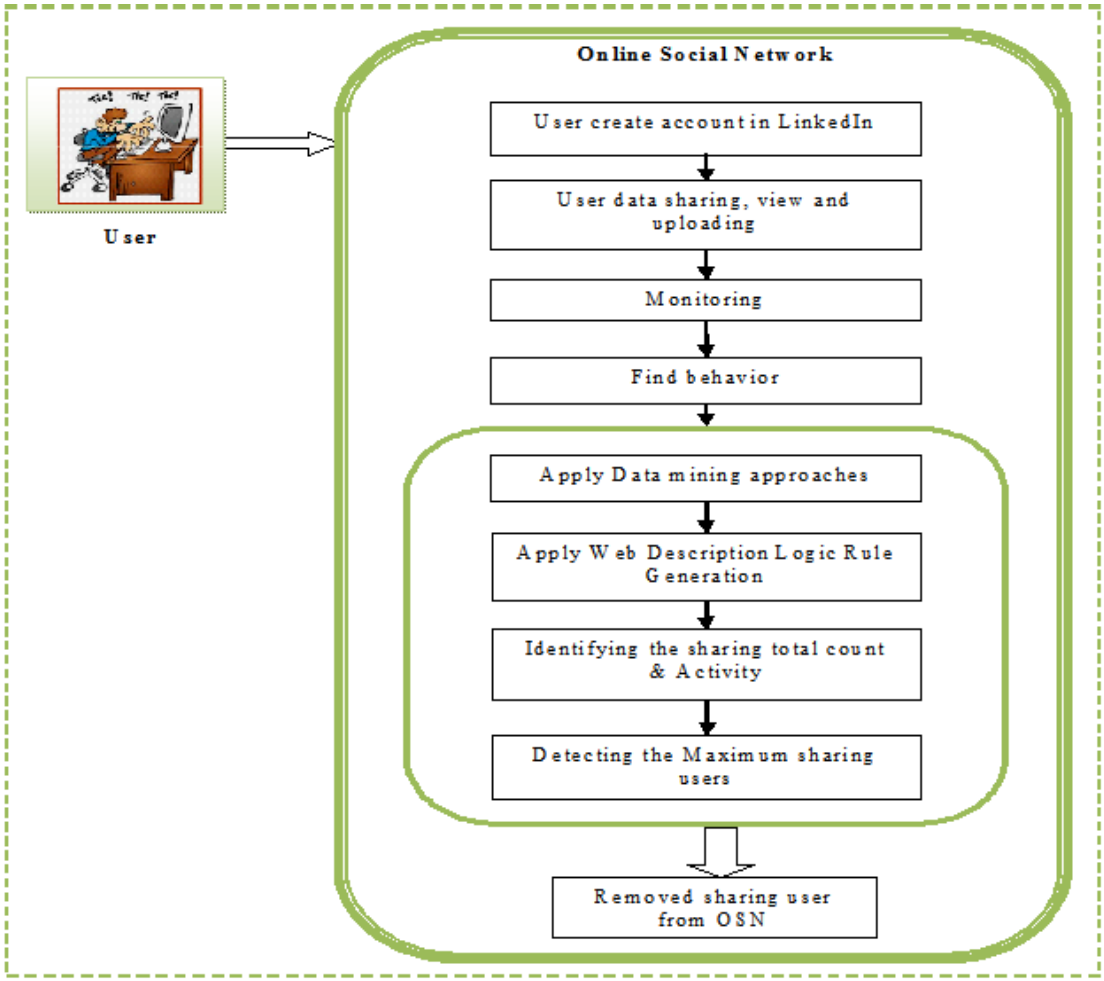

Figure 1: Overall Proposed Architecture

3.2 SOCIAL NETWORK CREATION

GUI is one user type in LinkedIn which lets users interact with other users using graphical icons as well as visual indicators like secondary notation, in contrast to the typed command labeling, or text navigation, or text-oriented interfaces. It becomes possible to have three user types in this

manner like owner of an account, TPS, and other users. It is then possible for the owner to upload privacy data into the system; public information / personal information that the owner shares are stored in third-party for uploading different kinds of data and different images such as the face images, natural images, general information, public information, personal information, and other information or images. The data that is uploaded may be of any size or any type. Users who upload electronic information like images on their home pages might want to also share the same only with their mutual friends – this can be partially satisfied by OSNs by making use of privacy settings. It becomes possible to view the electronic data shared in the OSN publically.

3.3 UPLOADING INFORMATION

Millions of world-wide users have become used to the online social networks (OSNs) that are now a part of their lives. Users are empowered to build explicit networks which are representing their social relationships while often sharing huge volume of personal data for their own advantage. The primary phase of a sharing system is the acquisition of image and personal data. It is possible for uploading different kinds of data and different images such as the face images, natural images, general information, public information, personal information, and other information or images. The data that is uploaded may be of any size or any type. Users who upload electronic information like images on their home pages might want to also share the same only with their mutual friends – this can be partially satisfied by OSNs by making use of privacy settings. It becomes possible to view the electronic data shared in the OSN publically.

3.4 PRIVACY SETTINGS

The information / images of each user is categorized first into privacy policy. Then each image's privacy policy may be classified and examined for exactly predicting the policy. Hence, adopt the 2-stage method regarding policy recommendation rather than applying a common single-stage data mining method for mining together both the policies and image features. The two-phased method lets the system engage the primary phase for classifying whether policy is with or without privacy. During the second phase, if it becomes possible to set without the privacy means, then details of user list may be preferred.

3.5 PROTECTED SYSTEM MONITORING

It can set protection or block the system for avoiding third-party attacks without the image owners' knowledge. This particular module could be used for setting images with privacy. In case a user is set with privacy settings, then every user may be considered third parties. According to such setting, it will be possible for unauthorized users to view the image only, and they cannot use it. Eventually, a system of hardware control such as keyboard controls and mouse controls can be provided. Print screen controls and mouse code are extracted; for providing coding implementation and to disable coding to be false settings.

3.6 MONITORING

Vulnerability user behavior as well as qualities related to LinkedIn can be monitored here. It will be then possible for LinkedIn users to both publically and privately send messages; options for this are built in. Users are also enabled to share their post with others. Users are enabled to search other users' public posts and profiles. Users may send and accept the friend requests also, in this module. Building the social behavior profile pertaining to individual users is empowered by properly combining the corresponding social behavioral measures. After that, we may describe application of social behavior profiles in detecting the malicious accounts and differentiating the users.

3.7 FIND BEHAVIOUR

User social conducts can be classified on the OSN into the two types, namely, introversive behaviors and extroversive behaviors. Extroversive behaviors include sending messages and uploading photos, resulting in the visible imprints of users to either one or even more other users. Contrarily, introversive behaviors like searching the message inbox and browsing profiles of other users however are not found to produce noticeable impacts on other users. The way in which any user normally interacts with friends online is reflected directly by the extroversive behaviors; hence they prove vital for the characterization of users' social conducts. Introversive behaviors form the major part of users OSN actions. So, the introversive behavior model forms the vital part of users' online social behavior qualities.

3.8 WEB DESCRIPTION LOGIC RULE GENERATION ALGORITHM

It has been presumed that users having several friends normally will be paying more attention toward building new friendship affinity in future and these users' preferences are impacted by their implicit and explicit friends' preferences. Web Description Logic Rule Generation Algorithm has been implemented in this work to find and analyze attackers and vulnerabilities. In this study to spot the most optimal unit set of aspects for detecting vulnerable friend in LinkedIn. By this algorithm vulnerability users are identified based on the sharing threshold values and Extroversive Behaviors to remove if he/she exceeds threshold limit and unwanted activity user.

ALGORITHM:

Step: 1 Start

Step: 2 Let consider set of S user in community S_n

$S = \{s_1 , s_2 , \ldots . s_n\}$

Input $I = \{I_1, I_2, ,,,,, I_n\}$

Step: 3 Assing sharing thershold to 0 for every users s,

for i = 1 to s_n do

u[i] = 0;

end for

Step: 4 Extroversive behaviors = $\sum_i^n S_i$ i=1, 2……n

Introversive behaviors = $\sum_j^n S_j$j=1,2……n

Step: 5 User Identify creations

For each $S_i \& S_j \in$ OSN Do

$I_n \rightarrow$ Regular Activity;

$S_i \& S_j \leftarrow I_n$;

end for

Step: 6 if any privacy vulnerable happens

Step: 7 check every user u[i] sharing percentage per day
thershold = 20;
MisActivity[S_i&S_j]= 30;
Step: 8 Find vulnerableusers in OSN;
for i = 1 to S_n do
if sharing count (n[i]) ≥ threshold && S_i,S_j≥ MisActivity) then{
For each(S_j≥ MisActivity) do{
n[i] S_j,→ Sharing user
Remove → S_i,S_j;
}
Remove (n[i])
else
n[i] → Non Sharing user
Continue (n[i], S_i,S_j)
end if
}
end for
}

3.9 ADVANTAGES

The system suggested here is like a tool that detects vulnerable user as well as application.
This provides protection to users' accounts from possible hackers.
This paper is a minor contribution in the specific area for offering protection to users and bringing down the volume of malicious actions.
It is give output result Maximum accuracy.
Minimum time

4. RESULT AND DISCUSSION

Today, online social networks (OSNs) are becoming very popular interactive medium for sharing, communicating, and spreading voluminous data. Continuous and daily communication leads to the exchange of many kinds of information like image, text, video, audio, and application. LinkedIn happens to be one among these social networking applications. In this work, It is used similar measurements of attributes that are created by applying expand measure over all the attributes by using bogus project dataset. We have suggested a method that spots vulnerability user in LinkedIn social network, the suggested method is based on determination of effective aspects for the process of detection. In OSN privacy vulnerability is main problem from users because some user spread other user privacy information. For avoiding this particular issue in OSN, In this paper have devolved Web Description Logic Rule Generation algorithm for discovering malicious users within the community or the group. Using this algorithm, minimizing the privacy vulnerable is made possible in OSN. Our proposed method is capable of providing better results in OSN to find vulnerable users when issues of privacy occur. We have gathered attributes from various researches; they have been spotted through extensive examination at the primary phase, and then these aspects have been properly weighted.

TABLE 1: COMPARISON CLASSIFICATION TECHNIQUES

S.No	Techniques	Accuracy (%)	Time (Ms)
1	Random forest	93.44	33:78
2	Decision tree	95.74	28.10
3.	Web Description Logic Rule Generation	97.77	17.12

In table 1, it has been proven that comparison of performance into identifying vulnerable user profile processing is used in the classification methods for comparing with many machine learning systems like Decision tree1. Feizy. R et, al,. , Web Description Logic Rule Generation, and Random forest. It has been assessed from this work that Web Description Logic Rule Generation (WDLRG) methods produce improved output compared to other present methods.

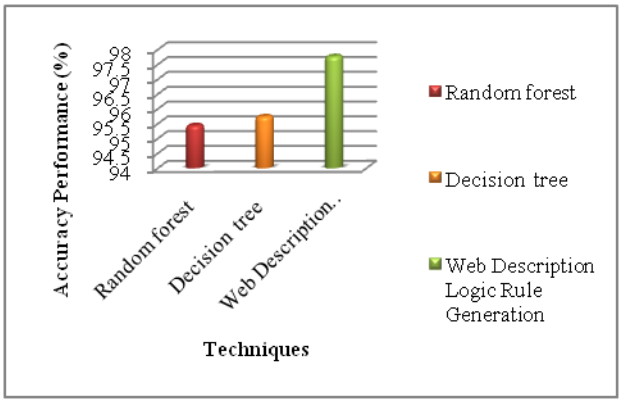

FIGURE 2: COMPARISON OF CLASSIFICATION TECHNIQUES

As given in bar chart, in the comparison of Graph performance in detecting the vulnerable user profile processing that is used in the classification methods for comparing with many machine learning approaches are Decision tree(DT), Random forest(RF) and Web Description Logic Rule Generation(WDLRG). It has been assessed from this research that Web Description Logic Rule Generation (WDLRG) algorithm produce improved output when compared with the other present methods.

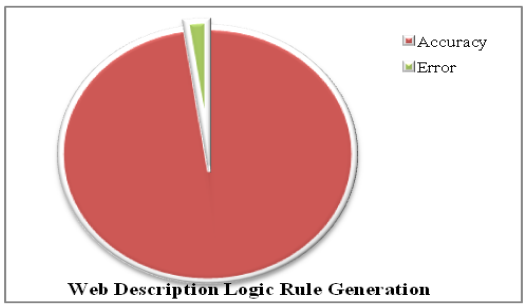

FIGURE 3: ACCURACY RESULT

As seen in pie chart, error performance to detect vulnerable user profile processing method output on the online social network shows an accuracy of 97.33% while error works out to 2.77%.

SIMULATION RESULTS

Experiments were conceded out based on the following configuration: Windows 7, Intel Pentium (R), CPU G2020 and processer speed 2.90 GHz respectively. The required software configuration is mentioned below, Operating System→Windows 7, Front End→JAVA, Back End→MYSQL. The datasets have been collected from the LinkedIn pages and the development is carried out in the JAVA environment. The implemented by whole dataset maintained in MYSQL and also getting from output results stored in MYSQL.

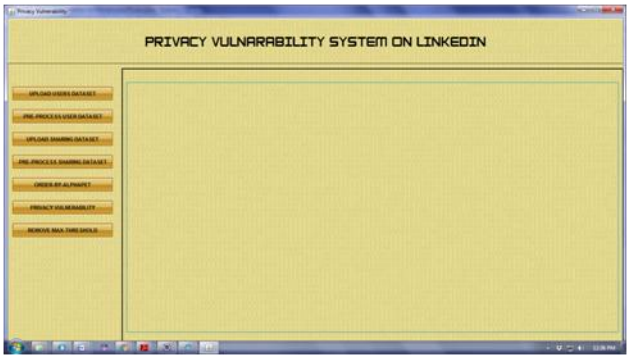

FIGURE 4: GUI PAGES

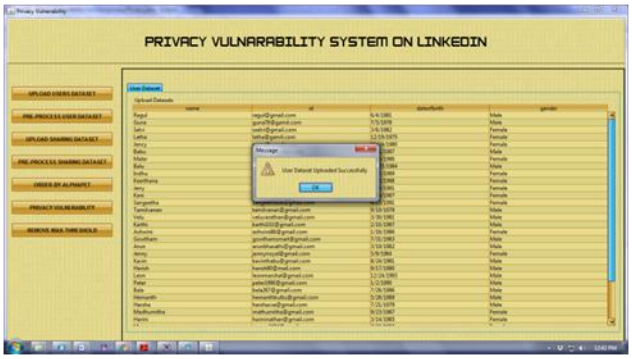

FIGURE 5: UPLOADED DATASETS

Fig 5, shows data set collection information from face book page. Data set are collected from different pages from one group to achieve find vulnerable users in this list.

FIGURE 6: PREPROCESSED DATASETS

Figure 6 shows user profile information of each user in that group. User information are gathered and created as dataset. This dataset is preprocessed for remove null and unwanted information from input dataset.

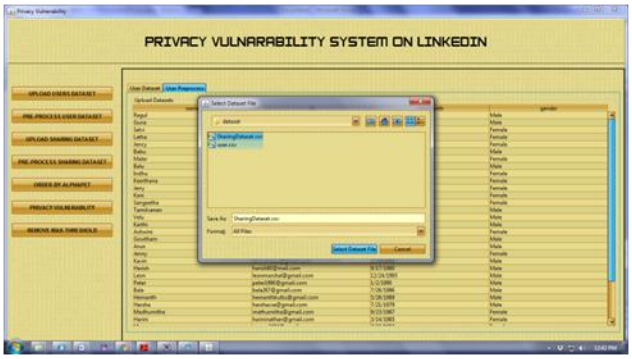

FIGURE 7: USERS SHARED DATASETS

Figure 7 explains user sharing information on group. This sharing information is collected and preprocessed for avoid unnecessary think. This sharing information is collected based on each user sharing posts.

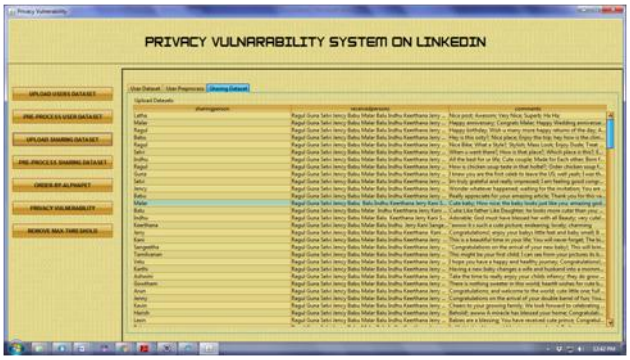

FIGURE 8: SHARED DATA PREPROCESSED DATASETS

Figure 8 shows user profile information of each user in that group. User information are gathered and created as dataset. This dataset is preprocessed for remove null and unwanted information from input dataset.

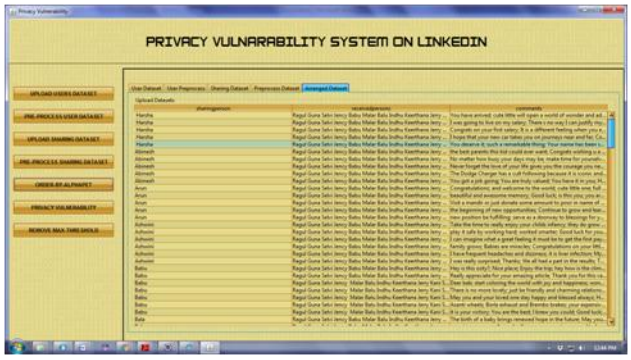

FIGURE 9: ARRANGED DATASETS

Figure 9 shows arranged dataset of sharing information. This sharing information is shuffled when upload into system. Then it will be arranged based on alphabetic order to identify each user sharing information.

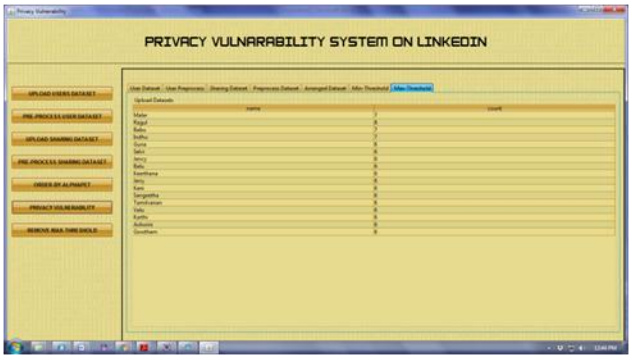

FIGURE 10: MIN AND MAX THRESHOLD DATA

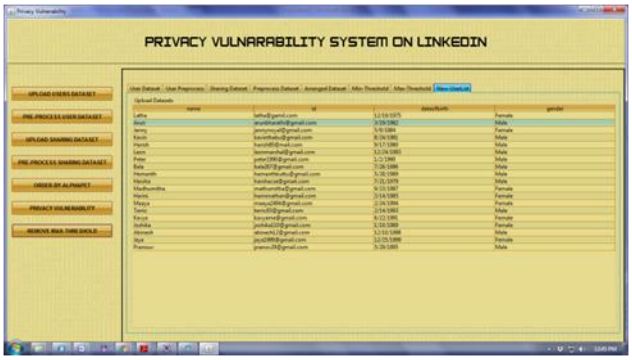

FIGURE 11: REMOVING THE USER

Figure 10 shows minimum and maximum threshold of each user sharing count. This minimum and maximum threshold count is used for identify maximum sharing person in this group. Then fig, 11 shows the user information after finding user threshold count. Users are removed from user profile list based on threshold count.

5. CONCLUSION

Online social networks now have attained the status of being the basic way of sharing and communication among people; they also are causes of concern regarding security and privacy. Social networks are extensively being used by students only in a rather small way toward educational purpose. A very large number of people are using social networks primarily for being in contact with those individuals who they know, and normally tend to divulge a lot of personal data. In this work, to have suggested a system to discover vulnerable user accounts on LinkedIn social network. The suggested method has been formed on determination of the effective features regarding the process of detection. Attributes from various researches have been gathered. They were identified via extensive evaluation as the first phase; then the aspects were weighted. Various tests have been conducted for reaching the minimum group of attributes, while perceiving the most optimal accuracy result. The paper suggests web description rule generation algorithm for removing effectively the malicious users out of the LinkedIn social network.

REFERENCES

[1] Fire, M., Kagan, D., Elyashar, A., &Elovici, Y. (2014). Friend or foe? Fake profile identification in online social networks. Social Network Analysis and Mining, 4(1), 1-23.

[2] Wald, R., Khoshgoftaar, T. M., Napolitano, A., & Sumner, C. (2013). "Predicting susceptibility to social bots on twitter". In Information Reuse and Integration (IRI), 2013 IEEE 14th International Conference on (pp. 6-13). IEEE.

[3] Probst, F., Grosswiele, D. K. L., &Pfleger, D. K. R. (2013). Who will lead and who will follow: Identifying Influential Users in Online Social Networks. Business& Information Systems Engineering, 5(3), 179-193.

[4] S. Mahmood(2013), "Online Social Networks: Privacy Threats and Defenses," in Security and Privacy Preserving in Social Networks, pp. 47–71, Springer.

[5] Gaurav Parsewar, Yogesh Dalvi, Lalit Kothwade(2015), "Detection of Malicious Application on Online Social Network" IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2.

[6] [M. Egele, G. Stringing, C. Kruegel, and G. Vigna,(2013), "COMPA: Detecting compromised accounts on social networks," in Proc. Symp. New. Disturb. Syst. Secur. (NDSS), San Diego, CA, USA.

[7] Adikari, Shalinda and Dutta, Kaushik, (2014), "IDENTIFYING FAKE PROFILES IN LINKEDIN" PACIS 2014 Proceedings. 278. http://aisel.aisnet.org/pacis2014/278

[8] Gaurav Parsewar, Yogesh Dalvi, Lalit Kothwade,(2015) "Detection of Malicious Application on Online Social Network" IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 9.

[9] Hsieh, C.-J., et al. (2013). "Organizational overlap on social networks and its applications." Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee.

[10] Shuliang WANG, Hanning YUAN,(2013), "Spatial Data Mining in the Context of Big Data", IEEE International Conference on Parallel and Distributed System.

[11] M.R. Khayyambashi, and F.S. Rizzi, (2013). "An Approach for Detecting Profile Cloning in Online Social Networks", 7th International Conference on e-Commerce in Developing Countries with Focus on e-Security (ECDC), IEEE.

[12] Thomas K,McCoy D, Grier C, et al. (2013) "Trafficking fraudulent accounts: The role of the underground market in Twitter spam and abuse." In: USENIX Security, pp.195–210.
[13] Jiang M, Cui P, Beutel, A, et al(2014). "Detecting suspicious following behavior in multimillion-node social networks". In: Proceedings of the 23rd International Conference on World Wide Web, pp. 305–306. ACM.
[14] Clark, EM, Williams JR, Jones CA., et al. (2016),"Sifting robotic from organic text: a natural language approach for detecting automation on Twitter." Journal of Computational Science 16: 1–7.
[15] Xiao C, Freeman DM and Hwa T.(2015)" Detecting clusters of fake accounts in online social networks". In: Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security, pp. 91–101. ACM.
[16] G.Narasimha Murthy, M. Eranna,(2017)" Detection of Behavioral Abnormality of Compromised Accounts in OSN'S", International Journal of Innovative Research in Science, Engineering and Technology", Vol. 6, Issue 10.