International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET

Composite Behavioral Modeling for Identity Theft Detection in Online Social Networks

[1]A Hima Bindu,[2]Sheelam Sriya Reddy, [3]Gurrala Sai Sri

[1]Assistant Professor, Department of CSE, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India

[2,3]Student, Department of CSE, Bhoj Reddy Engineering College for Women, Hyderabad, Telangana, India

ARTICLEINFO

Article History :

Accepted: 01 July 2023
Published: 10 July 2023

Publication Issue :
Volume 10, Issue 4
July-August-2023

Page Number :
73-78

ABSTRACT

Recent trends show that the popularity of Social Networks (SNs) has been increasing rapidly. From daily communication sites to online communities, an average person's daily life has become dependent on these online networks. Additionally, the number of people using at least one of the social networks has increased drastically over the years. It is estimated that by the end of the year 2020, one-third of the world's population will have social accounts. Hence, user privacy protection has gained wide acclaim in the research community. It has also become evident that protection should be provided to these networks from unwanted intruders. In this dissertation, we consider data privacy on online social networks at the network level and the user level. The network-level privacy helps us to prevent information leakage to third-party users like advertisers. To achieve such privacy, we propose various schemes that combine the privacy of all the elements of a social network: node, edge, and attribute privacy by clustering the users based on their attribute similarity. We combine the concepts of k-anonymity and l-diversity to achieve user privacy. To provide user-level privacy, we consider the scenario of mobile social networks as the user location privacy is the much-compromised problem. We provide a distributed solution where users in an area come together to achieve their desired privacy constraints. We also consider the mobility of the user and the network to provide much better results.

Keywords : Data privacy, Social networks, User data, K-anonymity, Mobile social networks, Location based services, Dynamic clustering

I. INTRODUCTION

Usage of Online social Networks is ever-increasing. In 2017, 71 percent of all internet users were using some kind of social networking site. It is expected that by the end of 2019, there will be 2.77 billion social network users. This constitutes 2/3rd of the world's population [8]. Hence, it is clear that the vast majority of the population uses OSN, and their profiles are online. As the initial OSN sites were introduced, connections in OSN were only based on real-life friends. As the market grew and people started to explore, OSNs now offers a wide range of friend-finding systems. We can now find friends based on the place we live, common interests, common friends, and so on. Hence, the public profile of a user has become an important part of one's profile. Without correct information about a user, it is difficult to find things or people that he would be interested in. Hence, users need to have a profile that contains personal information, including their date of birth to a recent place that they have visited. Also, this public profile contains a lot of sensitive information like age, gender, profession, current area, and so on, that includes his / her photo(s). As discussed before, public profiles are a common feature of any OSN. Due to this fact, a huge amount of personal data is vulnerable to attacks. One of the biggest scandals of 2018 is the Cambridge Analytica. According to CNBC, The Guardian, The Observer, and the New York Times, it is suspected that two of the biggest social media sites Facebook and Analytica allegedly used user's personal information for the 2016 presidential election in the USA. It is claimed that 50 million user profiles have been mined to send them personalized messages that contributed to the election results.

II. RELATED WORK

Social Networks (SNs) have attracted millions of people worldwide from the time they were introduced. In today's world, they have become an indispensable part of every human's life. There are many ways in which a social network is defined. According to Webster [1], a social network is nothing but a network of people. However, it is clear from our usage of social networks these days that it is so much beyond connecting people. Social networks are often used for online communication and hence, can be referred to as Online Social Networks (OSNs). They are now not just a means of communication between people. The social network industry has grown and spread its business into every part of our daily life. More and more social networks are emerging every day for various requirements, and every social network is a booming business model. Some of the early social media sites include Friendster, Six Degrees, Orkut, and so on. Orkut became one of the premier OSN sites in Brazil and India during 2007 [2]. Another popular instant messaging system, QQ, was in widespread use in China is introduced in 1999 [3]. However, today's market is hugely dominated by Facebook and Instagram that combinedly has about 290 million users [4]. The main purpose of an OSN site is to provide web services that will allow users to perform the below actions:

1. Have a public profile to present to other users.
2. Have friends/connections to communicate with them
3. Search for new friends based on their various attributes like common friends, interests,

and so on

The above functionalities are the basic functionalities based on which a social network is initially developed. Hence, we categorize all the social networks into one/more of the following categories based on the works done by [5] [6] [7].

1. Personal networks. OSNs like Facebook, Friendster and MySpace are some examples of this category. This category of networks focuses on creating a detailed user's profile. Hence they allow many attributes for a single user.
2. Status update networks. Twitter is the best example of this category. These networks mainly focus on

posting a status update. These updates might sometimes include a place or person.

3. Shared-interest networks. This category of networks focuses on bringing people with similar interests together. Dating apps like Tinder and professional apps like LinkedIn are examples of this category.

4. Neighborhood Exploring networks. These networks highly focus on user's current and exact location. Based on user's location, he will be able to share information, media files and interact with neighborhood people. Tinder is also an example of this category. With more and more increase in the demand for various functionalities, social networks now do not belong to a single category. They are trying to expand and provide more features to please the users. For example, Facebook was initially a personal network, and now it can be considered as all the four categories mentioned above.

In today's world, social networks have developed into multi-goal applications. They are not just used to connect people but for business, finding local businesses, dating, job search, and many more. Also, as mobile phones boom started in the early 2000s, every major business website tried to enter the mobile application world. That formed the basis for the development of Mobile Social Networks (MSNs). With the introduction of 4G and LTE, internet speeds have skyrocketed, and the mobile applications started using real-time data of the mobile user. All the social network providers have their mobile application that supports various functionalities. Hence, a social network is now divided into OSN and MSN applications. Although MSN is a part of OSN, there are more complications when privacy is to be provided for mobile users compared to a stationary user. In the following section, let us examine the privacy threats on OSN.

III. PROPOSED SYSTEM

As mobile devices' usage has tremendously increased, Mobile social Networks (MSNs) have become an integral part of our lives. According to the recent survey by Statista [6], there are over 61% of users in North America use MSNs. It is expected that the number of users who use MSNs reaches around 2.46 billion in 2017 and 3.02 billion in 2021. This number is almost one-third of the current earth's population. By integrating various technologies and applications, social network providers market their products to reach a wide range of users. Their applications are compatible with PDAs, smart phones, and many more. The primary advantage of such MSN applications is that they provide various services like LBS, virtual reality, online dating, and so on. Among them, location-based queries are the most widely used service. The Online Social Network (OSN) has become a host for various business advertisements. These business models include hotels, restaurants, vacation spots, and sports. Users of OSN visiting these business models tend to review them, which in turn helps other users to visit them. Hence the business owners maintain very detailed information of their business, such as location, menu, phone numbers, and address. As the mobile internet has exploded with its increasing speed, humongous mobile applications have been developed in the past years. Similarly, all the OSN providers have their mobile-based applications designed to provide a more convenient and faster way to access their accounts on the go. MSN applications offer various advantages, including LBS, that help users to search for nearby business models. LBSs have been gaining considerable popularity by the rapid advances in positioning technologies, e.g., Global Position System (GPS), and the development of modern smart devices with data communication capabilities. LBS query is the most common usage of any mobile user. Users query for nearby restaurants, the distance between their position and another place, and so on. A typical example of such a service is: Alice would like to query for nearby restaurants. For such queries, all we need is the exact location but not the user profile information. If the user and his location are identified by the attacker who can snoop the network, then important

information can be leaked. For example, Alice visits a cancer hospital regularly and goes to a pharmacy at a particular location whenever she visits the hospital. Imagine she has enabled LBS to any of her online social networking applications. Then, the attacker can infer that Alice or a family member of hers has cancer, and also they live in the neighborhood of the pharmacy she visits. This is a piece of crucial information that should not be leaked. This attack can be seen in Fig. 1.

While mobile users enjoy LBS, they have to provide their real locations to the LSP, which poses a severe threat to their privacy. It is to be observed that the users might not always want to reveal their location information to others. The key to address these concerns lies in the preservation of the users' privacy when efficiently providing correct query results. However, some LBS does consider user profile. For example, if a person searches for a nearby restaurant, the result contains all the restaurants in the current area but sorted according to the user history or interest. If the user preference is Asian cuisine, then the results are sorted accordingly, and so on. Hence, accurate results for LBS is only possible if we have exact location information and correct user profile. However, if we provide those to the LBs query, then the privacy of the user is not maintained. Let us denote a time series of social graphs as G0,G1,. . . .,GT . For each temporal graph Gt = (V, E, Lt), the set of vertices is V and the set of edges is E. Lt is the location set of all

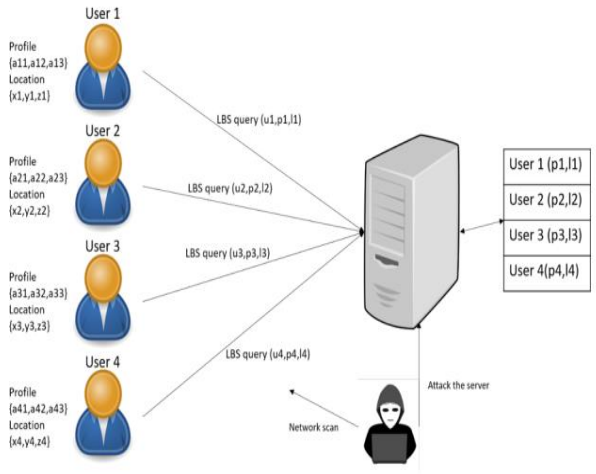

Figure1. Problems of existing LBS query with user privacy

the users at time 't'. For our theoretical analysis, we focus on undirected graphs where all the |Et | edges are symmetric, i.e. (i, j) ∈ E if and only if (j, i) ∈ E. In each temporal graph G, vertices denote the users and edges denote the connection between them. Given a user 'u' with location, 'l' wants to search for LBS with users of similar interests. Let us consider that there are 'n' users in the location radius 'r' each with 'A' attributes. We obfuscate user 'u' based on the equi-cardinal clustering and send out the obfuscated user details 'o' with its generalized attributes.

U = u1, u2, u3, ..., un

ui = a1, a2, a3, . . . , aA

For a given 'k', this paper aims to obfuscate 'u' into 'o' in such a way that:

• There are at least 'k' users with the same attributes as 'o'

• Minimize information loss

IV. RESULTS AND ANALYSIS

This paper reports the results of a 21-day field experiment about the use of two types of nudging mechanisms to influence teenagers' posting privacy behavior in the social network platform PESEDIA. Nudge mechanisms proposed in this paper did not limit participants' ability to share information in the social network. Instead, they encouraged the participants to reflect on their potential audience that may have access to the information. In general, previous soft-paternalism approaches not only in the context of social networks state that these mechanisms make users reflect and become more aware of their decisions, avoiding risky behaviors [6, 2, 3].

Initially, we thought that the "learning curve" of a new social network platform such as PESEDIA would influence the users' privacy behaviors. However, after the analysis of the behavior of users without mechanisms during the period of the experiment, we found that there was not a significant difference in

their posting privacy behavior between the initial days of the experiment and the last days. There is significant evidence that users' privacy behavior for posting actions changed when the nudging mechanisms were activated. Independently of the mechanism used (i.e., picture or number nudge), when the nudging mechanisms were activated, the number of messages published with a private policy (i.e., only me, collections, or friends) was higher than the number of messages with a public policy. Therefore, this change could be driven by the nudges. Although users seem to publish with a more restrictive privacy policy, we noticed that most of them used friends or private policies without considering collections (i.e., a personalized subset of friends). This could be because the use of this policy in PESEDIA requires the manual creation of the collection or because it is a concept that is not present in the social network platforms that they are used to, and, therefore, they do not initially consider it as a possible option. Previous studies already showed the importance of nudges for increasing users' awareness about privacy and, thus, modify their behaviors. In this paper, we focused our experiments on teenagers, who are usually less concerned about privacy risk [1]. Although the effect of nudging mechanisms was appreciated, it is expected that more visible behavioral changes can appear if the experiment was extended in time [8].

Previous works that proposed the use of different types of nudge mechanisms do not pay attention to the differences between them on users behavior [2]. In this experimental study, we analyzed whether there is a significant difference between the effects on the privacy posting behavior of teenagers that had the Picture Nudge or the Number Nudge activated. The results revealed that there are no significant differences between mechanisms. This could be because the teenagers were focused mainly on the highlighted text about the risk level than on other details such as the profile pictures of users that may see the publication or the number of users that may see the publication.

In the literature, we cannot find studies that sharply measure the effect of some type of nudge to be more beneficial in terms of changing the posting behavior. However, some authors such as [4] and [8] state that the design of nudges that are more tailored to users would cause these nudges to be more effective. This would require aspects such as not receiving alerts about information that is already known or designing personalized nudges according to what is more effective for each specific user. This can be viewed as a limitation of our proposal that can be explored in future works. With regard to the perception of users about the nudges, the majority of teenagers considered nudges to be useful mechanisms to preserve their privacy in posting. This follows the results obtained by Wang et. al. [3] where the users that were involved in a similar experiment with nudges in social networks mentioned that nudges could be more useful for people without experience in social networks (i.e., teenagers). Al though the majority of the participants perceived nudges as beneficial, some of them considered them as irritating, and this is considered as a disadvantage towards the effective implementation of privacy nudges [4]. Wang et. al. [3] suggested that this behaviour can be associated to the profile of the publications (personal or not), but there is not any clear study that demonstrate this fact. In line with what is stated above, future research line should consider the design of more personalized nudges that really show information that is really valued by the specific user.

V. CONCLUSION

Online social networks are a powerful tool for getting a range of social benefits that traditional (offline) communication cannot offer. However, social networks are still not the most secure tools. Specifically, a lot of privacy issues have been reported about the users' actions on sharing information and privacy decisions. Therefore, advances in improving

social networks, privacy mechanisms, and, in turn, the users' privacy decisions are required.

VI. REFERENCES

[1]. (2017). Children and parents: Media use and attitudes report.

[2]. Abar, S., Theodoropoulos, G. K., Lemarinier, P., and O'Hare, G. M. (2017). Agent based modelling and simulation tools: A review of the state-of-art software. Com puter Science Review, 24(Supplement C), 13 – 33.

[3]. Abbasi, A., Chung, K. S. K., and Hossain, L. (2012). Egocentric analysis of co authorship network structure, position and performance. Information Processing & Management, 48(4), 671–679.

[4]. Abril, D., Navarro-Arribas, G., and Torra, V. (2011). On the declassification of confidential documents. In International Conference on Modeling Decisions for Artificial Intelligence, pages 235–246. Springer.

[5]. Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509–514.

[6]. Acquisti, A., Taylor, C., and Wagman, L. (2016). The economics of privacy. Journal of Economic Literature, 54(2), 442–92.

[7]. Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., et al. (2017). Nudges for privacy and security: Understanding and assisting users' choices online. ACM Computing Surveys, 50, 44.

[8]. Akhtar, N. (2014). Social network analysis tools. In Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on, pages 388–392.IEEE.

[9]. Al-Rahmi, W. M., Alias, N., Othman, M. S., Marin, V. I., and Tur, G. (2018). A model of factors affecting learning performance through the use of social media in malaysian higher education. Computers & Education, 121, 59–72.

[10]. Albert, D. and Steinberg, L. (2011). Judgment and decision making in adolescence. Journal of Research on Adolescence, 21(1), 211–224.

[11]. Alemany, J., del Val, E., Alberola, J., and García-Fornes, A. (2018). Estimation of privacy risk through centrality metrics. Future Generation Computer Systems, 82, 63–76.

[12]. Alemany, J., del Val, E., Alberola, J., and García-Fornes, A. (2019a). Enhanc ing the privacy risk awareness of teenagers in online social networks through soft paternalism mechanisms. International Journal of Human-Computer Studies.

[13]. Alemany, J., Del Val, E., Alberola, J. M., and Garćia-Fornes, A. (2019b). Metrics for privacy assessment when sharing information in online social networks. IEEE Access, 7, 143631–143645.

[14]. Alemany, J., Del Val, E., and García-Fornes, A. (2020). Empowering users regard ing the sensitivity of their data in social networks through nudge mechanisms. In Proceedings of the 53rd Hawaii International Conference on System Sciences, pages 2539– 2548.

[15]. Alemany Bordera, J. (2016). PESEDIA. Red social para concienciar en privacidad. Mas ter's thesis, Universitat Politècnica de València, Valencia, Spain.

Cite this article as :

A Hima Bindu, Sheelam Sriya Reddy, Gurrala Sai Sri, "Composite Behavioral Modeling for Identity Theft Detection in Online Social Networks", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 4, pp. 73-78, July-August 2023.

Journal URL : https://ijsrset.com/IJSRSET23102127