International Journal of Safety and Security Engineering
Vol. 12, No. 4, August, 2022, pp. 507-518
Journal homepage: http://iieta.org/journals/ijsse

Comprehensive Analysis of Privacy Attacks in Online Social Network: Security Issues and Challenges

Sandip A. Kahate[1*], Atul D. Raut[2]

[1] Research Centre P. G. Computer Science Department, SGBA University, Amravati 444601, Maharashtra, India
[2] CSE Department, P. R. P. P. College of Engineering. and Management, Amravati 444601, Maharashtra, India

Corresponding Author Email: sandipk_ioe@bkc.met.edu

https://doi.org/10.18280/ijsse.120412

Received: 18 July 2022
Accepted: 20 August 2022

Keywords:
online social network, security, privacy issues, malicious user, machine learning, decentralized technique

ABSTRACT

Nowadays, users value their privacy of information is more than money but the Online Social Network is creating new platforms for cybercrime to intimidate innocent users of countries because of the privacy pitfalls in the present traditional centralized architecture like Facebook, WhatsApp, and Twitter, Instagram, and many more. Third-party apps and malicious attackers breach innocent users' private information, especially adolescent users for their personal purpose. Many Asian countries have no special data protection laws like a European. The main objectives of this research paper are to study the different types of privacy attacks for data confidentiality in the Online Social networks with analyzed mitigation techniques and discussed future proposed work on how to control and detect the user's private information from unreliable people. Even then, few privacy destructions are unresolved!

1. INTRODUCTION

Online social networks are becoming tremendously popular among young as well as old people just too virtually interconnect with each other for the purpose of exchanging information and entertainment while supporting various new apps that are launching day by day.

Michael Fire et al. survey the OSNs, those apps which stand as a demanding rank in the youth of society, such as Facebook, Google+, Twitter, Instagram, etc., as listed in Figure 1.

Figure 1. Online social network

It has more demanding apps in society, primarily from young people [1]. Worldwide, there are over 2.9 billion monthly active Facebook users as of the first quarter of 2022 which is a 13 percent increase year over year. YouTube has 2.5 billion monthly active accounts, while WhatsApp has 2 billion. Instagram, the fourth-most popular photo sharing app, with over 1.4 billion users and 436 million monthly active Twitter users globally. Global internet users spent an average of 147 minutes each day on social media, so, in 2022, over 4.26 billion individuals used social media globally, with that figure predicted to climb to over six billion by 2027 [2].

Through his or her profile, an OSN user creates his or her own identity in the social network, which is accessible to their friends in a transitive manner. OSN also has the ability to create links between different users. He/she can form these connections with various users known as "friends," "mutual friends," and "friends-of-friends," and even accepts friend requests from strangers who may turn out to be good friends. If both the users are successfully connected with each other then it is considered as a neighbor. Chewae et al. examine how the high demand for and regular use of OSN causes security and privacy issues in cybercrime. Non-secure private information in OSNs will result in the open entry of attacks for susceptibilities or malicious users for destructive intentions, especially by teenagers [3]. Nowadays, in the area of OSNs, many researchers are working on security and privacy aspects that make the OSN systems more appropriate to society at large in the future [3-6]. Privacy setting also provided by many OSNs to allow users to avoid other users' access. As survival is one of the important criteria, intervention of advertising agencies, political parties during election prevent the providers to breach information (theft identity) to third parties.

The main purpose behind all of these in the picture is to show the centralized infrastructure of OSN under the control of a single administrator. Hence, users have no options without believing in the OSN provider to protect all their confidential data. Even though the user does not know whether their confidential data is actually protected from attackers who may breach and theft from the provider's server or not, Therefore, OSN users' privacy is an important issue [7]. Recently, maximum changes have been made in OSN to increase the privacy and security features, but there are still some limitations which cannot be totally overcome related to safe sharing of confidential data but it is possible by using machine or deep learning models to detect these OSN attacks, we have illustrated in proposed worked.

In future approach, decentralized trustable immutable blockchain network with cryptography light weight OSN attacks resistant algorithm and smart contract techniques to prevents the attacks and improve the data confidentiality of users.

In this proposed approach we have evaluated on different sizes of dataset (consisting of Twitter tweets and breach datasets) obtained from Kaggle and other online sources. Experiments for the suggested approach were conducted on three datasets for complete result analysis using machine learning classifiers to detect data breaches, cyberbullying, and cyberstalking assaults in online social networks. We investigate the performance of these classifiers using the performance measures accuracy, precision, recall, and F1 score. We have also analyzed the positive, negative, and neutral words. It is based on the NLTK library's sentiments analyzer. Finally, we compared our proposed method's results to past similar studies.

Section 2 deals with motivation and objectives. Related work in the state of art to describe all existing work with the limitations presented in Section 3. Section 4 illustrates the working of an online network with a sequence diagram of creating login credentials. Section V describes the current scenario of online social networking along with an elaboration of all attacks on OSN. Existing mitigation techniques are available in Section 6, result and analysis using Machine Learning Algorithms model predicted in Section 7 and conclusions are in Section 8. Section 9 points to future work, which includes a description of a possible plan by many researchers to implement new OSN structures in the future to tackle these problems. One such plan talks about decentralized techniques in distributed online social networks (DOSNs) with a block-chain based framework [8, 9].

2. MOTIVATION

Before launching OSNs, information exchange could happen in digital media in one-to-one fashion and was less risky than sending it in one-to-many fashion.

Nowadays, people around the globe use social network for different purposes. They upload their locations [10], videos, images etc. on the social media. However, they even don't know (especially adolescents user) that they are victim to very serious matters that can happen due to their uploading post and compromise their privacy. These incidents can potentially put them and everyone around them in potential danger. Many serious incidents have occurred and have been reported as a result of careless postings or the sharing of confidential or private information on the unsecured infrastructure of OSNs. Today's huge class of sufferers are teenagers who widely use social networking sites [11]. In this scenario, awareness about the people who are viewing personal data like marital status, age, school/college, photos and many more shared information plays vital role with a point of OSNs user is concern at the cost of connecting with people for fun and entertainment.

Objectives of online social network:

- To describe critical evaluation of existing malicious detection system.
- To perform comprehensive threat analysis and identify attack modeling using potential security technique verification tools.
- To analyze and predict the performance of existing machine learning for adolescent predator detection, comparison and perform gap analysis.

3. STATE OF ART

Hosseiny et al. [12] presented and evaluated the results for detection of DOS attacks using a decision tree machine learning classifier model. Grasshopper and ant colony optimization algorithms with PSO and genetics techniques have increased accuracy and speed in attack detection using different features of the NSK-KDD dataset.

They have worked on one machine learning model that detected a single DOS attack. But there has been scope for researchers to work on different cybersecurity attack detection using some more machine learning models for comparing the best evaluation metrics results.

As we know social media is the hot topic nowadays, but all analysis done in traditional way. In future research required enhanced method to identify the identity theft crimes into sites, Upadhyay et al. [13] has performed operations on some algorithms for cyber bullying detection: Adult Image Detection Algorithm, Irrelevant Posts Detection Algorithm, NLP Algorithm.

Proposed cyberbullying system is not able to detect user that number of times to post a thing and even does not detect if he/she has posted something by mistake. Need to work on foreign words vocabulary.

Ebrahimi et al. [14] elaborated NLP Algorithm: some categories of feelings are defined using algorithms, like when we should reject or ignore the person if he or she has posted how many times, intentionally or by mistake. However, additional work could be done to elaborate on vocabulary variation with relevant features used in chat rooms. But the accuracy of deep learning using CNN is not satisfied.

Ngejane et al. [15] suggested different types of ML base algorithms with datasets as Support Vector Machines (SVM) accurately classify pedophiles versus cybersex conversations k-Nearest Neighbors (k-NN) to cluster only pedophile conversations, Convolutional Neural Network (CNN) auto detection of sexual raider identification.

But still more work to be done in terms of feature engineering deep learning models with semi supervised may be improve the accuracy on CNN and emerged LSTM and have been in Natural Language Processing and categorization of text may be influence on new ML models.

Persia and D'Auria [16] elaborated operator solution: In this involves Authentication Mechanisms, Security and privacy settings. Commercial solution: FB phishing protector, MacAfee Social protections, Defensio, Nortan safe web. Vulnerabilities are only studied and presented in this paper.

Al-Garadi et al. [17] analyzed to detect aggressive behavior on social websites by using ML attitudes. Outmoded manually Machine Learning supervised algorithms absence the capability to handle cyberbullying. Research is required to developed effective and accurate model for cyberbullying detection. Hence Deep Learning has a new technique to attract many researchers.

Praveena and Smys [18] in this paper "Privacy preservation by k-anonymization of weighted social networks", of nodes against attacks where the opponent has information but for anonymization detect solution by TDR in classification. OSNs user data privacy has been developed by k-anonymity and L-diversity but this techniques loss the user information and developed one time released network data but not dynamic.

Soumya and Revathy [19] implemented and described E Classic threads solution Phishing: Guard against spams, Spamming: Use of CAPTCHA and bot un- friendly Stalking: Profile privacy, Avoid strangers report stalkers. Above all solutions are only used for detecting the threats in online social sites but not remove permanently.

In their paper, Pamungkas et al. [20] evaluated results from different machine learning algorithms. In the paper, they used some datasets for evaluation with one hot encoder for preprocessing and an autoencoder technique to detect intrusion in the system. In their proposed system, they have worked on reducing the feature dimensions and softmax techniques. When we reduce the feature dimension of data, it then significantly increases the accuracy, but sometimes we lose the important feature from the data. There is also a need to reduce the size of data to improve the accuracy as well as many feature selection algorithms to enhance the detection of intrusion.

4. WORKING OF ONLINE SOCIAL NETWORKS

Dürr et al. [21] Facebook has been studied and shown to be a hugely popular tool on social networking sites among a large class of young people for sharing private information with one another. Different age groups People also use Facebook to connect with communities of people for different purposes, like professional and academic work. Another social demand app, GitHub, is most popular among IT professionals for sharing their work, programming logic, coding, and algorithm information with each other in a community [21].

A study by Gangopadhyay and Dhar [22] states that LinkedIn can provide rights to users to share information among their group members only if another group member can view their full profile, can request for all resource material within group members (known friends), so it is a little bit more secure and maintains privacy than other existing social networking sites. One can easily connect with same interest person on the LinkedIn network through sending a request communication or sharing of information is possible in LinkedIn social network only if request is accepted [22].

Figure 2 shows user profile creation and login sequence of OSN site that demonstrate how to create a profile and the login page on Social Network site. Identifications is mandatory for a user to create an account and login credentials for validation and verification purpose. After creating account, a user can accomplish all operations like send a friends request, videos, social network, photos, and post through OSN using login credentials [23].

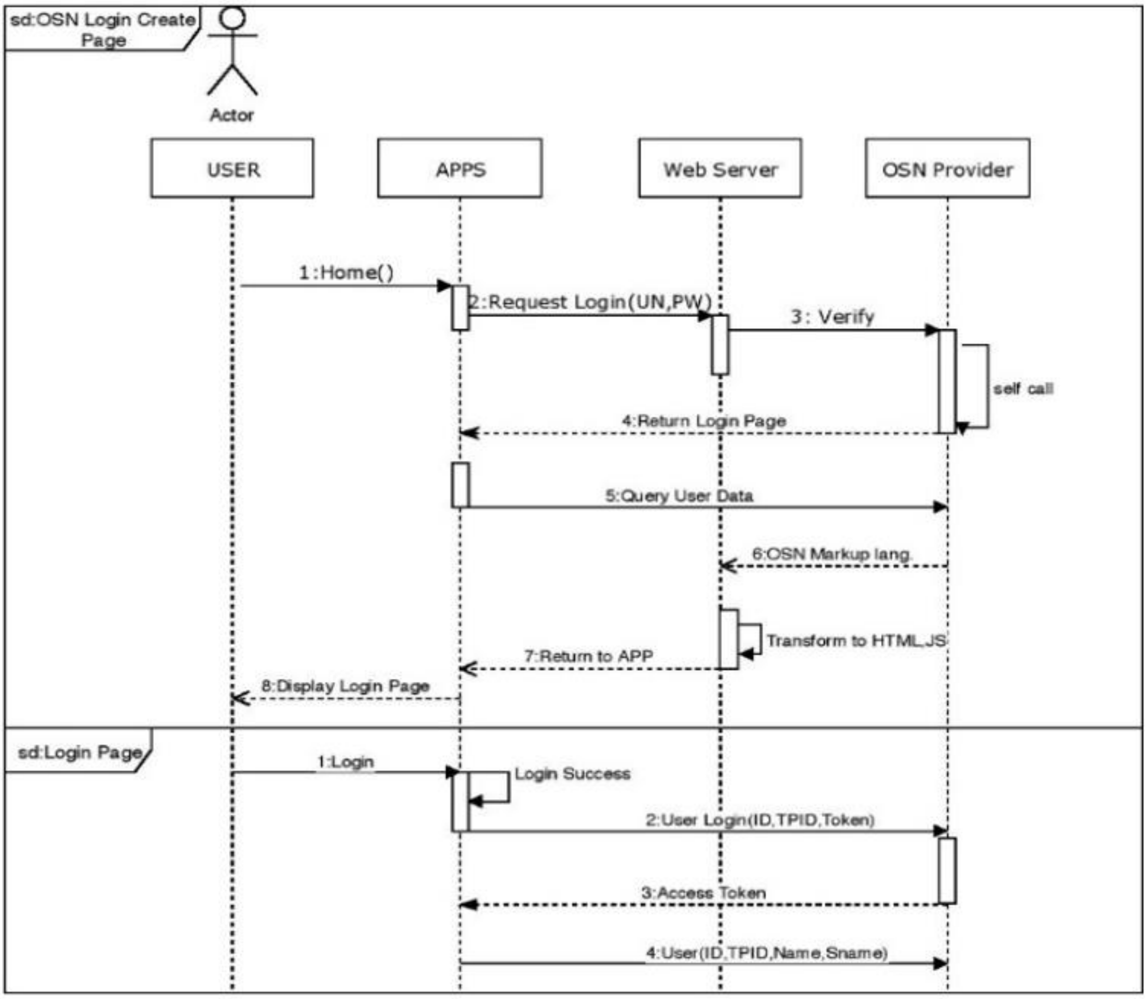

Figure 2. OSN user profile creation and login

5. HOW SAFE ONLINE SOCIAL NETWORK

Alrubaian et al. [24] stated that, due to the strong, instantaneous, and cheaper communication channels of online social media, the world has now demanded a worldwide site for comment on different events. OSNs plays vital role in connecting with family members, friends circle, members of similar interest, community, commercial groups. Event organizers taking the benefits of all social network sites for their advertising, marketing, business people for improving business reputations, political parties for campaigning at the election time because of privacy setting now available in the almost all OSNs [24].

Unfortunately, maximum users (laymen in IT knowledge) do not have knowledge of the privacy policies of the OSNs they are using. Many of them are unaware of the privacy control feature and default privacy settings. Few OSNs like Facebook track their users' profiles and without a profile, users of different sites and third-party game sites. It also captures bio-metric thumb impression or facial data without user's explicit "opt-in" consent and offers advertiser for their use of interest. So, user's confidentiality can be at threat.

Definitely, OSNs tools/apps will make room for new hackers and criminals to perform fraudulent and undesired activities and attacking to inject malware web link, viruses, phishing, spamming, etc., become a result of identity theft.

Huang et al. presented that it has now become clear that social networking sites are so easy to access by adolescents who are not security conscious. Each OSN has its own rules and dataset for sharing data within group users only. So, it is significant issues that how we should secure our confidential information on the online social network from intruders [25].

5.1 Privacy issues in OSN

We know that today's vast user crowd is available on social networks, but on the other hand, intruders are attracted to and access users' private information by developing various methods of malevolent intent to acquire and analyze such personal data. OSN can perform anticipated behavior is to be sure entity, but when users sharing their important personal information as a wealth, at the same time, he/she should also consider what undesired activities might take place.

We will analysis the role of OSN, in terms of its privacy, and possible attacks on users' privacy [26]. The word privacy has been defined with two aspects. One from personal privacy (physical privacy) to information privacy, each with own definition. Information privacy is relevant to the web on privacy, as defined below in the IITF wording that Kang uses [26]. Second Information Privacy is "personally claim to control the terms under which personal information identifiable to the individual is acquired, disclosed or used" [27].

Openly, the challenge is for both users as well as admin in many times, users are unknown about breaching of their information. Information can be breach and theft by an anonymous person for illegal use. In next subsection we deliberate the main privacy and security issues associated with online social network and explores specific attacks with models which are alert to adolescent user and for secure sharing of their personal information.

5.2 Attack modeling

The different attacks modeling is listed and illustrated as below, in which user's security and privacy are in threats. These attacks are mainly divided into three categories shows in Figure 3. Martin M, et al listed all the categories of attack modeling. The first category contains classic attack which is included with Malware, Phishing, and Spammer, XSS, and Internet fraud. Modern attacks given in the second category are completely different in the OSN environment, which uses the OSN structure to threaten user privacy and security. These are click-jacking, de-anonymization, location, and information.

The next two threads social engineering and identity theft are described in subsections 1) and 2). The third and last category includes bullying specifically targeting children those are use social networks. These are Online Predator, Cyber-Bullying and Cyber Grooming. Cyber-Bullying and Cyber-Grooming are deeply explored in subsection 3) and 4) [28].

1) Social-Engineering Attacks: Albladi and Weir stated that social engineering is an essential aspect of information security. Multiple victims have been targeted by attackers by using different enhanced techniques [29]. Even with increased concern of risk, in this research work on social engineering illustrate the attacks which are possible by reverse social engineering [30], phishing [31], or direct attacks, when users account hacks to theft, confidential private data may be used for other purpose [32], or user profile link with across the multiple social network [33] might be possible of social engineering attacks on their personal or organization data. Social engineering attackers usually trap the victims.

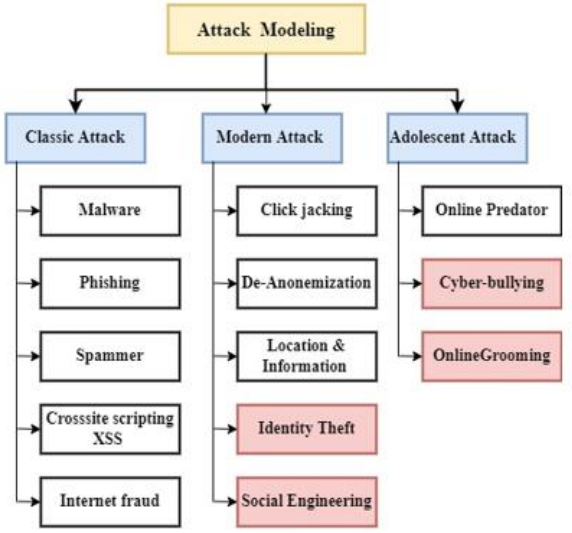

Figure 3. Online social network attack modeling

2) Identity Theft: C. Chipurici detected and studied and said about Identity Theft that it is an attack by using the fake credentials of the attacks which is a more influence on the targeted victim. Social networking sites like Instagram, and Facebook have essential apps for each life those who just simple basic knowledge about the use of these tools. Everybody well known about this platform that it has become a background for intruders to theft information [34]. Though the concept of Identity Theft has been old, only change the techniques. Those methods are like Phishing, Identity spoofing and hacking to acquiring personal in-formation by intruder stated by Hoelscher [33].

Figure 4 shows a sympathetic of scam in sequence diagram. In which intruders attempt to track, identify, and access confidential data, such as login credentials, by impersonating other people. In the first step, the attacker usually performs some malicious techniques to crack the OSN adolescent user's password and requests the OSN server as a legitimate user for the picture which is depicted in the sequence diagram. In step two, the next server responds and displays all of the images and confidential information. Then the attacker captures all the victim's confidential information many times from the OSN server, as illustrated in consecutive steps 2.1–2.3.

Finally, in step 3, a malicious user informs the victim that "I have all your information" and starts harassment/blackmailing the victim. When the victim realizes that the attacker has stolen (breached) all information, he/she has inform the OSN administrator (Server).

Farhoud et al. [35] illustrated first inject ill-behaved links into fake social sites that capture user credentials and otherwise consider the victim as a friend in their group by sending a mysterious post, which the user naively accepts the friend request, and then malicious users breach identities of victims' accounts from the server of social network. This happens mainly due to usage of same password of all apps. Now the malicious user can get access, send porn images for harassment to victims on their social accounts, and also hack the personal contact number, which is linked with the Unique Identification Authority of India (UID), bank account, credit or debit card, or even handle the user's social network account for illegal activities.

3) Cyber-Bullying: Al-Garadi et al. explained the Caber-stalking or Cyber-bully is hot topic for today's researchers that cyberbully as using electronic communication method to intimidate user mostly teenagers are victim [36]. Cyberbully means such as harassing, writing, making hateful or aggressive posts for insulting a victim [37, 38]. Cyberbully is considered as dangerous and fast- spreading aggressive behavior which can be easily committed. Intimidates only require electronic devices like cell phone or computer connected to the network to perform troublesome without bullying the users [39]. Hinduja and Patchin [40] went on to say that nowadays, school students are completely reliant on social networking for academic work or any other reason with their friends and relatives, which exacerbates cyber-bullying issues.

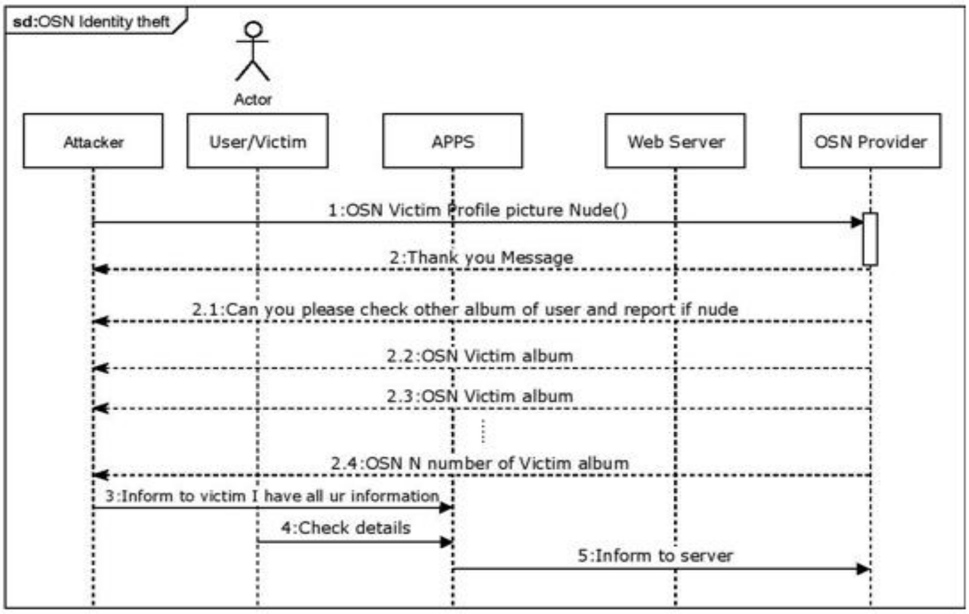
sd:OSN Identity theft
Actor
Attacker
User/Victim
APPS
Web Server
OSN Provider
1:OSN Victim Profile picture Nude()
2:Thank you Message
2.1:Can you please check other album of user and report if nude
2.2:OSN Victim album
2.3:OSN Victim album
2.4:OSN N number of Victim album
3:Inform to victim I have all ur information
4:Check details
5:Inform to server

Figure 4. OSN identity theft attack

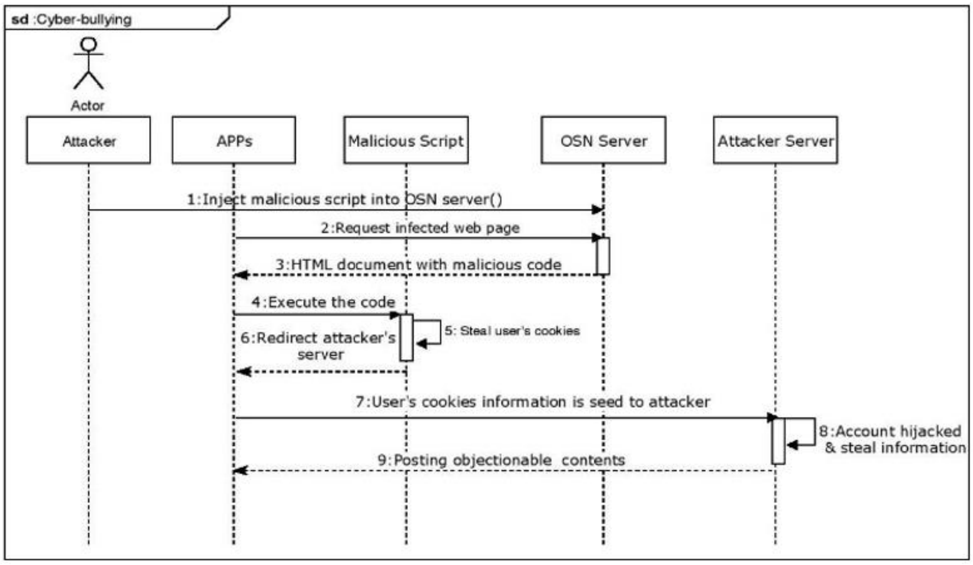
sd :Cyber-bullying
Actor
Attacker
APPs
Malicious Script
OSN Server
Attacker Server
1:Inject malicious script into OSN server()
2:Request infected web page
3:HTML document with malicious code
4:Execute the code
5: Steal user's cookies
6:Redirect attacker's server
7:User's cookies information is seed to attacker
8:Account hijacked & steal information
9:Posting objectionable contents

Figure 5. OSN cyber-bullying attack

Because online social network tools have become such a marvelous part of their lives, but some predator take the advantages through this technology for malicious use toward others [41]. For example, in Figure 5. Explained teenagers created their account and post private room images, videos, audios, location and created pages on social networks for just fun and entertainments where privacy is expected, but attackers take benefits to sending hurtful text to server and capture their private information for harassment using smart devices.

Attackers also have recorded unauthorized videos of their peers and black-mail them to upload or post for the world to see, likes, tag, and discuss. We always alert teenagers that cyberbully means when suspected person "frequently sends some text, videos posts something online about another person that they don't like or accept their friend request."

4) Online Cyber-Grooming or Cyber-Stalking: Online grooming [15]: is clear by Harms [41], as "a communication process by which a perpetrator applies affinity seeking strategies, while simultaneously engaging in sexual desensitization and information acquisition about targeted victims in order to develop relationships that result in need fulfillment". Like a physical sexual person. This terms also use for sexual predator to describe for such people and may be interchangeably use this word. Girouard [42] described about teenagers those are daily user of social networking with similar approach. Author further said that one among six kids of ages 13-19 years come up to online for sexually activities.

Young predators are snooping and will engage in online discussion that they would keep secret all about things from real world. Due to poor decision taking from teenagers by blindly believing on unknown stranger of similar interest (by giving likes to their post) leads to high risks like open online discussion about sexual activity.

6. EXISTING MITIGATION TECHNIQUES

Irshad and Soomro [43] stated some techniques for identity theft in social networks that extreme use of OSNs has improved the privacy and security settings but still has some limitations to give the chance for attackers to attack and breach the safe sharing of information and cause economic losses. But Instagram, GitHub, LinkedIn, and other social networking sites are enhancing the privacy settings to protect users' privacy and control the maximum security issues like breaching confidential data [43]. According to Wright et al. "the zone of internet information security is technologically advanced and continues to advance in response to new bullying" [44].

Ananthula et al. illustrated some key points that might be help to protect the online social network users from attackers and request to don't posted your important information like personal contact number, current or vacations location and photos on social sites may be converted in a Cybercrime. Always, think that Internet as a public. But don't forget that all information is stored on a centralized server that can be hacked even through your privacy settings [45]. Always beware from strangers instead of claiming every time because they could have theft your identity to commit crime and always confirmed before using or installing third-party apps from social network sites. If you want to track your online activities, then you should install malware [46, 47]. For strong passwords, use six to eight characters, including special characters and numbers, and change them every few months. Use anti malware software and also install antivirus software to tackle threats. Every parent monitored their kids during the use of any social networking sites because they are unaware about wise techniques of online security and how to report suspected person with security threat.

In this research paper, we are taking the work that has already been illustrated by other authors into deliberation and some only point out an overall idea to implement new techniques in the future for teenagers to protect them from online grooming and cyberbullying. Upadhyay et al. implemented as an Adult Image Detection Algorithm: In this work, the author analyzed the disciplined use of images and videos. The traditional text analytics framework consists of text pre-processing [48], NLP Algorithm: some categories of feelings are defined using algorithms, like when we should reject or ignore the person if he or she has posted how many times, intentionally or by mistake. However, additional work could be done to elaborate on vocabulary variation with relevant features used in chat rooms [49].

Ebrahimi et al. studied Dataset: They used PAN-12 for SVM (Support Vector Machine) with semi-supervised. This model is trained on destructively and tested on unlabeled data. Their tentative results show an accuracy of 98%, then again proceed with deep learning techniques and using same dataset experiment with CNN (Convolutional Neural Network) for sexual predator identification with auto detection method shows the result only 80% [50] Ngejane et al. further expanded previous limitations on CNN models to improve accuracy. By experimenting with new ML models, LSTM (Long-Short Term Memory) has been developed and demonstrated to be capable of text categorization and natural language processing [15, 16].

7. RESULTS AND ANALYSIS

As we all know, social networking is a trendy issue these days, yet many academics perform their research in an old fashioned manner using outdated approaches. The study now requires the creation of a more comprehensive technique for identifying social network attacks on websites.

This section presents the findings of the analysis on current datasets before adopting the system model.

In the future, final results will be derived using real-time statistics lawfully scraped from social networking web-sites such as Twitter, Facebook, and others. However, before adopting the model, we extensively study the data to get insight into the dataset by conducting analysis on three current OSN assaults that we addressed in Section V: identity theft cyber bullying, and cyber grooming or cyber stalking. This section described the performance parameters used for evaluation with python language tool.

A. Performance evaluation and algorithms:

The performance parameters are used in this analysis are:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$\text{Precision}=\frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{3}$$

$$\text{F1 Score}= \frac{2*(Precision*Recall)}{(Precision+Recall)} \quad (4)$$

where:

- TP : True Positive
- TN : True Negative
- FP: False Positive
- FN : False Negative

B. Performing results on following algorithms:

This system uses machine learning algorithms to establish a correlation between the positive and negative tweets on Twitter. Actually, there are three types of machine learning algorithms: supervised, unsupervised, and reinforcement learning. We employed supervised learning techniques to map the essential Characteristics of the target variables utilizing training data that included vectors and their related outputs. Among the widely used supervised learning algorithms chosen for implementation are:

- SGDClassifier
- Logistic Regression.
- Decision Tree.
- Random Forest.
- LinearSVC

C. Experiment performance:

1) Identity theft: The collected dataset is pre-processed and then performing analysis on the identity theft threats by plotting Pie chart. The dataset contains 298 samples that have been categorical features converted in five different values. Insights in Figure 6. presents a positioning of the most breaches of data. It reports around 62.4% of the data breaches by hacking method [20].

Time complexity of classification algorithms for data breaches:

Time complexity refers to the ability of algorithms to perform efficiently on input sizes of dataset attributes with smaller capacity in order to improve results. So we performed a time complexity study on a dataset of 298 sample as an input (n) and computed the results using certain machine learning models or algorithms. According to Figure 7. the best training time complexity is generated by KNeighborsClassifier with 0.018 seconds and the worst is created by KMeans with 0.480 seconds. Next we observed the best and worst prediction time complexity of giving the input size (n) feature datasets to the different classifier models or algorithms. The MultinomialNB classifier algorithm produced the best prediction result with a time complexity of 0.086 seconds, while the KMeans algorithm produced the worst with a time complexity of 0.480 seconds.

***2) Cyber bullying:* Time complexity of classification algorithms for cyber bullying dataset:** The primary goal of any analysis is to maximize accuracy while minimizing time complexity. The accuracy and time complexity of classification algorithms are measured and compared. Accuracy informs us how successfully classification is achieved, whereas time complexity tells us how long it takes to examine the data. In the Figure 8 graph, x axis represents algorithms and y axis time complexity. The major goal was to evaluate the dataset's performance and observed the pattern of change in accuracy and time complexity as the dataset's levels were gradually reduced. The best prediction time complexity we saw with LogisticRegression is 0.057 seconds, and the worst prediction time complexity we observed with RandomForestClassifier is 0.811 seconds for the first level of the dataset, and as we moved on to decrease the dataset level, it created an accuracy while decreasing the time complexity of the dataset [51]. We observed that all the algorithms followed the same trend. As we gradually decrease the size of the datasets.

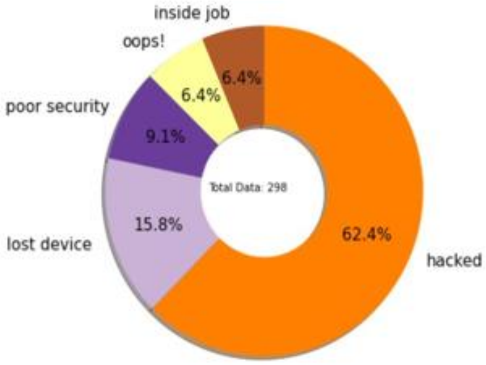

Figure 6. Distribution of identity theft

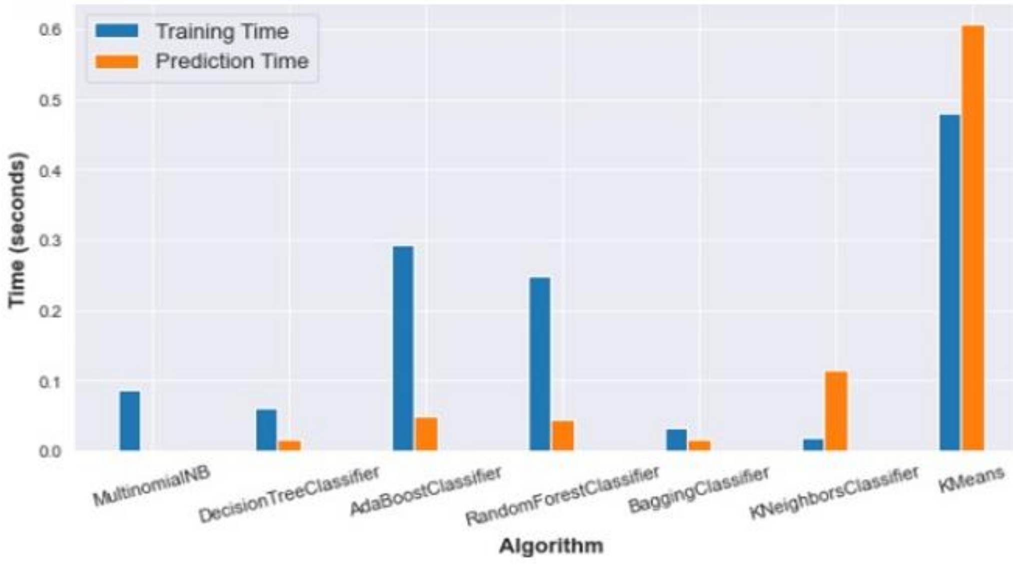

Figure 7. Breaches time complexity

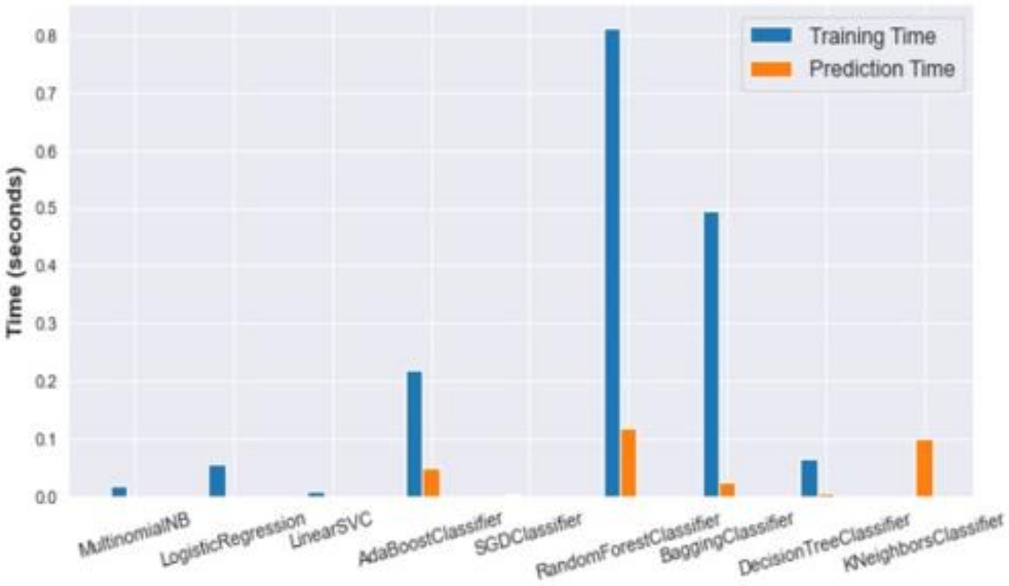

Figure 8. Cyber bullying time complexity

Classification summary of algorithms on cyber bullying dataset:

In this study, we have acquired Github data sets and pre-processed them. The dataset is multiclass, and the recall and accuracy of each class were independently determined using the confusion matrix of each approach. The datasets comprise 1063 samples divided into 797 training and 266 testing purposes on 75%-25%. Scaling, Vectorization, transformation and pipeline were used in the matrix process. Then send it to be evaluated against several algorithm models. We have observed the predicted values. Table 1 summarizes the average values of four metrics for each model.

Table 1. Average accuracy, F1-score, precision and recall algorithms

Algorithms	Average Accuracy	F1-score	Precision	Averag Recall
SGDC	0.6955	0.6197	0.6168	0.6226
Logist_Reg	0.7143	0.6238	0.6562	0.5943
DecisionT.	**0.7218**	**0.6408**	**0.6600**	0.6226
RF	0.6617	0.5545	0.5833	0.5283
LinearSVC	0.6992	0.6330	0.6161	**0.6509**

The algorithm with the highest F1 score is Decision Tree algorithm. Thus, decision tree algorithm has the highest performance in comparison to the other algorithms even if the accuracy, Precision and Recall score suggests otherwise. Average Accuracy, F1-score, Precision and Recall Algorithms Best performance by decisionTree given as follows: Avg Accuracy: 0.7218 F1- score: 0.6408 Precision: 0.6600 Recall: 0.6226.

3) Cyber grooming or cyber stalking: First and foremost, we employ a Python Word Cloud dependency that can rapidly display us the most often used words in Cyber Grooming related tweets. Our word cloud depicts the top 100 most frequently used words in tweets. We created a word cloud for the top 100 words, with the larger the word in the cloud, the more popular it is (you can change this parameter by altering the value for ’max words’). Using the NLTK tool package, we can identify whether all tweets from scraped datasets are positive, negative, or neutral [52].

- **Neutral words:** Some of the regular words like people, MKR, think, Known and etc. In general, we cannot tell if these tweets express positive or negative feelings about cyber grooming or stalking, therefore we may consider them neutral words, as seen in Figure 9.
- **Positive words:** Next, we look at the distribution of positive tweets. This analyzer examines the sentiments of a sentence according to whether it is positive, and it is based on the sentiments analyzer of the NLTK library. We can interpret the sentiment in the Figure 10. If a sentiment is positive, it could mean that it is not harmful to the Online Social Network (OSN) user and does not commit Cyber Crime. A negative sentiment could imply that it is harmful to adolescent social network victims while being supportive of grooming threads.

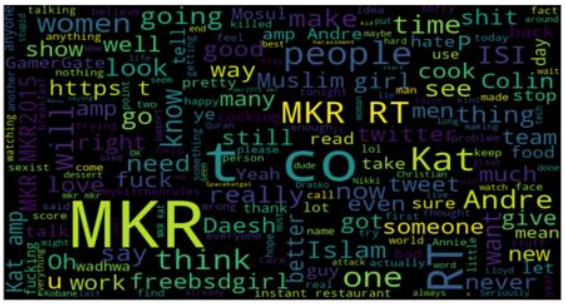

Figure 9. Neutral words

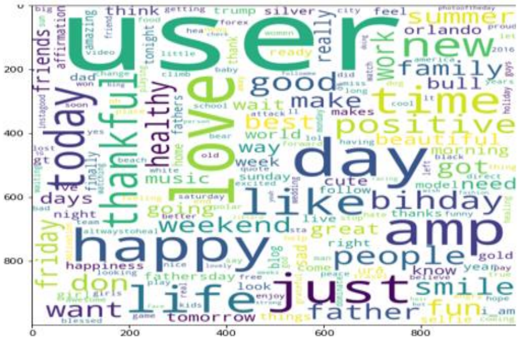

Figure 10. Positive words

Table 2. Average training accuracy, validation accuracy, f1-score

Algorithm	Training Accuracy	Validation Accuracy	F1 Score
RandomForest	0.99	0.95	0.61
LogisticRegression	0.985	0.94	0.59
DecisionTree	0.99	0.93	0.53

- **Negative words:** Figure 11 shows random tweets with negative sentiments. The tweets words like sexiest, hate, female, bitch etc.

This means that the tweets are SUPPOSED to be negative tweets. These types of tweets determine the emotional blackmail, hate, and harassment of adolescent users on Twitter, Facebook, or any other social site.

Classification summary of algorithms on cyber grooming dataset: Table 2 shows the assessed results on the GitHub dataset of three models created using the 31962 sample dataset divided in a 75:25 ratio for training and validation purposes. Furthermore the normalized confusion matrix derived from the test dataset yielded good results. By obtaining larger True Positives and True Negatives levels. Finally, the estimated training accuracy, validation, and f1-score using Eqns. (1)-(4) show that the trained RandomForestClassifier outperforms the projected performance outcomes.

D. Comparative outcomes with related results

Our evaluated outcomes of proposed work were compared with the past existing results of experiments performed by previous researchers [49-53].

The comparative results are shown in Table 3. In this, we have observed that our proposed ML classifier model for Decision Tree provided good average accuracy outcomes for cyber-bullying while Random Forest has the best test accuracy as compared to others classifier models.

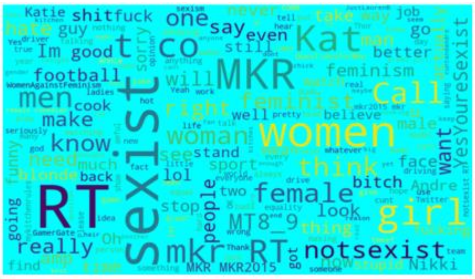

Figure 11. Negative words

Table 3. Comparative results with existing works

Paper	Approach	Dataset	Advantage / Disadvantage	Test Accuracy
[49]	Logistic Regression Support Vector Machine	BOW for Dataset	BOW used to achieve better Prediction Time	85% 82.7%
[50]	Logistic Regression Random Forest	Global Tweets datasets-37373	Did not investigated many feature extractions techniques.	90.57% 89.84%
[52]	Random Forest AdaBoost Classifier Support Vector Machine (SVM)	Datasets from Kaggle	TF-IDF used but not sufficient to completely detect cyberstalking cases.	84.43% 82.44% 87.41%
Our results on **Cyber-bullying** Dataset.	Decision Tree (DT) LogisticRegression	Twitter 1063 samples dataset	Evaluated on all ML models. We observed that Decision Tree classifier good worked on average Accuracy than others for **Cyber-Bullying** attack detection.	**72.18%(Avg)** **89%**
Cyber-Stalking Dataset.	DecisionTree Random Forest	Dataset-49159 from Repository	Outperforms the projected performance outcomes on **Cyber-Stalking** attack detection.	**92%** **95% (Best)**

8. CONCLUSION

In this research paper, we have proposed theft identity (Data Breach), cyber-bullying, cyber-stalking, or cyber-grooming detection classifiers. Investigated using Machine Learning (ML) models based on BOW, TF-IDF, and Word2Vec feature extraction of Natural Language Processing (NLP) have been used on global repository datasets.

In experiments, we have observed that 62.4% of the data breaches by the hacking method in identity theft as well as the best training time complexity are generated by KNeighborsClassifier with 0.018 seconds.

We analyzed prior studies in the realm of cyber-bullying and cyber-grooming using Twitter datasets and measured metric performance. Proposed approaches on several ML classifiers are implemented, and the results are compared.

According to this investigation, Decision Tree for Cyberbullying earned an average projected test accuracy of 72.18% when compared to other accuracy measures, while Random Forest for Cyber Stalking provides 95.0% validation (test) is greater malicious detection accuracy.

However, our current methodologies are insufficient to identify cyber-bullying and cyber-grooming harmful assaults on online social networks, particularly in real-time and automated detection.

Because feature extraction is difficult for whole text comments, we divided the negative (0) and positive (1) words using performing label and One-Hot-Encoding, thus we did not use many feature extraction methods in this research.

9. FUTURE WORK

In future, we plan to work on malicious attacks detection for more accuracy using NLP techniques and Deep Learning many-to-one classifier model and design and proposed light weight cryptography algorithm for data confidentiality and design new efficient techniques which is only used in cryptosystem for securing currency but not in OSN to protect our important personal information using Blockchain approach for attack prevention.

Usually, OSN users keep their sensitive profile information on a centralized system that is easily hacked or vulnerable by malicious attackers, so by using block-chain techniques, we can store our information on a decentralized immutable system to avoid third-party interference.

ACKNOWLEDGMENT

I am extremely grateful to my research supervisor Dr. A. D. Raut, P.R. Pote Patil College of Engineering and Management, Amravati and mentor Dr. D. N. Chaudhari, Jawaharlal Darda Institute of Technology and Engineering, Yavatmal. Affiliated to Sant Gadagebaba Amravati University,Amravati, for their genuine support. It was a great privilege and honor to work and study under their guidance. I would like to express my deep and sincere gratitude to my beloved Dr. M. U. Kharat, Professor and Head in Computer Engineering Department and Dr. S. V. Gumaste, head of IT Department at MET's Institute of Engineering, Bhujbal Knowledge City, Nashik, affiliated to Savitribai Phule Pune University, Pune, for providing invaluable guidance throughout preparing this paper. His dynamism, vision, sincerity and motivation have deeply inspired me.

I express my special thanks Dr. V. M. Thakare, Professor and Head, Computer Department of Sant Gadagebaba Amravati University, Amravati and Dr. P. N. Mahalle, Professor and Head in Artifial Intelligrnce and Data Science Department VIIT, Pune for their support to do this work. Finally, my thanks go to all the people who have supported me to complete this paper directly or indirectly.

REFERENCES

[1] Fire, M., Goldschmidt, R., Elovici, Y. (2014). Online social networks: Threats and solutions. IEEE Communications Surveys & Tutorials, 16(4): 2019-2036. https://doi.org/10.1109/COMST.2014.2321628

[2] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/, Published by s. Dixon, accessed on 22 Aug. 2022.

[3] Chewae, M., Hayikader, S., Hasan, M.H., Ibrahim, J. (2015). How much privacy we still have on social network. International Journal of Scientific and Research Publications, 5(1): 2250-315.

[4] Jiang, L., Zhang, X. (2019). BCOSN: A blockchain-based decentralized online social network. IEEE Transactions on Computational Social Systems, 6(6): 1454-1466. https://doi.org/10.1109/TCSS.2019.2941650

[5] Yan, Z., Feng, W., Wang, P. (2015). Anonymous authentication for trustworthy pervasive social networking. IEEE Transactions on Computational Social Systems, 2(3): 88-98. https://doi.org/10.1109/TCSS.2016.2519463

[6] Khater, S., Gračanin, D., Elmongui, H.G. (2017). Personalized recommendation for online social networks information: Personal preferences and location-based community trends. IEEE Transactions on Computational Social Systems, 4(3): 104-120. https://doi.org/10.1109/TCSS.2017.2720632

[7] Cho, J.H. (2018). Dynamics of uncertain and conflicting opinions in social networks. IEEE Transactions on Computational Social Systems, 5(2): 518-531. https://doi.org/10.1109/TCSS.2018.2826532

[8] Tamane, S., Solanki, V.K., Dey, N. (Eds.). (2017). Privacy and security policies in big data. IGI Global.

[9] Nilizadeh, S., Jahid, S., Mittal, P., Borisov, N., Kapadia, A. (2012). Cachet: A decentralized architecture for privacy preserving social networking with caching. In Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies, pp. 337-348. https://doi.org/10.1145/2413176.2413215

[10] Klukovich, E., Erdin, E., Gunes, M.H. (2016). Posn: A privacy preserving decentralized social network app for mobile devices. In 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 1426-1429. https://doi.org/10.1109/ASONAM.2016.7752436

[11] Yamin, M., Abi Sen, A.A. (2018). Improving privacy and security of user data in location based services. International Journal of Ambient Computing and Intelligence (IJACI), 9(1): 19-42. https://doi.org/10.4018/IJACI.2018010102

[12] Hosseiny, S.M., Rahmani, A.I., Derakhshan, M. (2020). Improve intrusion detection using grasshopper optimization algorithm and decision trees. International Journal of Safety and Security Engineering, 10(3): 359-364. https://doi.org/10.18280/ijsse.100307

[13] Upadhyay, A., Chaudhari, A., Arunesh, Ghale, S., Pawar, S.S. (2017). Detection and prevention measures for cyberbullying and online grooming. 2017 International Conference on Inventive Systems and Control (ICISC). https://doi.org/10.1109/ICISC.2017.8068605

[14] Ebrahimi, M., Suen, C.Y., Ormandjieva, O. (2016). Detecting predatory conversations in social media by deep convolutional neural networks. Digital Investigation, 18: 33-49. https://doi.org/10.1016/j.diin.2016.07.001

[15] Ngejane, C.H., Mabuza-Hocquet, G., Eloff, J.H., Lefophane, S. (2018). Mitigating online sexual grooming cybercrime on social media using machine learning: A desktop survey. In 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), pp. 1-6. https://doi.org/10.1109/ICABCD.2018.8465413

[16] Persia, F., D'Auria, D. (2017). A survey of online social networks: Challenges and opportunities. 2017 IEEE International Conference on Information Reuse and Integration (IRI). https://doi.org/10.1109/IRI.2017.74

[17] Al-Garadi, M.A., Hussain, M.R., Khan, N., Murtaza, G., Nweke, H.F., Ali, I., Gani, A. (2019). Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges. IEEE Access, 7: 70701-70718. https://doi.org/10.1109/ACCESS.2019.2918354

[18] Praveena, A., Smys, S. (2017). Prevention of inference attacks for private information in social networking sites. In 2017 International Conference on Inventive Systems and Control (ICISC), pp. 1-7. https://doi.org/10.1109/ICISC.2017.8068648

[19] Soumya, T.R., Revathy, S. (2018). Survey on threats in online social media. In 2018 International Conference on

Communication and Signal Processing (ICCSP), pp. 0077-0081. https://doi.org/10.1109/ICCSP.2018.8524200

[20] Pamungkas, I.G.A.K., Ahmad, T., Ijtihadie, R.M. (2022). Analysis of autoencoder compression performance in intrusion detection system. International Journal of Safety and Security Engineering, 12(3): 395-401. https://doi.org/10.18280/ijsse.120314

[21] Dürr, M., Werner, M., Maier, M. (2010). Re-socializing online social networks. In 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing, pp. 786-791. https://doi.org/10.1109/GreenCom-CPSCom.2010.18

[22] Gangopadhyay, S., Dhar, D. (2014). Social networking sites and privacy issues concerning youths. Global Media Journal: Indian Edition, 5(1).

[23] Beye, M., Jeckmans, A., Erkin, Z., Hartel, P., Lagendijk, R., Tang, Q. (2010). Literature overview-privacy in online social networks. Centre for Telematics and Information Technology, University of Twente.

[24] Alrubaian, M., Al-Qurishi, M., Alamri, A., Al-Rakhami, M., Hassan, M.M., Fortino, G. (2018). Credibility in online social networks: A survey. IEEE Access, 7: 2828-2855. https://doi.org/10.1109/ACCESS.2018.2886314

[25] Huang, Q., Singh, V.K., Atrey, P.K. (2018). On cyberbullying incidents and underlying online social relationships. Journal of Computational Social Science, 1(2): 241-260. https://doi.org/10.1007/s42001-018-0026-9

[26] Kang, J. (1997). Information privacy in cyberspace transactions. Stan- ford Law Review, 50(4): 1193-1294.

[27] Salunke, M.B., Mahalle, P.N., Dhotre, P.S. (2022). Comprehensive threat analysis and activity modelling of physical layer attacks in internet of things. Handbook on ICT in Developing Countries: Next Generation ICT Technologies, 237.

[28] Martin, M.C., Lam, M.S. (2008). Automatic Generation of XSS and SQL Injection Attacks with Goal-Directed Model Checking. In USENIX Security Symposium, pp. 31-44.

[29] Albladi, S.M., Weir, G.R.S. (2018). User Characteriics that Influence judgment of social engineering attacks in social networks. Human-Centric Computing and Information Sciences, 8: 5. https://doi.org/10.1186/s13673-018-0128-7

[30] Irani, D., Balduzzi, M., Balzarotti, D., Kirda, E., Pu, C. (2011). Reverse social engineering attacks in online social networks. In International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, pp, 55-74. https://doi.org/10.1007/978-3-642-22424-9_4

[31] Shindarev, N., Bagretsov, G., Abramov, M., Tulupyeva, T., Suvorova, A. (2017). Approach to identifying of employees profiles in websites of social networks aimed to analyze social engineering vulnerabilities. In International Conference on Intelligent Information Technologies for Industry, pp. 441-447. https://doi.org/10.1007/978-3-319-68321-8_45

[32] Luo, W., Liu, J., Liu, J., Fan, C. (2009). An analysis of security in social networks. In 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, pp. 648-651. https://doi.org/10.1109/DASC.2009.100

[33] Hoelscher, P. (2017). Retrieved from: http://resources.Infosecistitute.com/category/nterprise/phishing/the-phishing landscape/ phising-attacks-by demographic/social networks/gref, accessed on 12 August 2022.

[34] Chipurici, C. (2016). Retrieved from: https://heimdalsecurity.com/blog/howto-preventidentity-theft-20-steps/, accessed on 12 August 2022.

[35] Farhoud, N. (2016). Retrieved from: http://www.mirror.co.uk/news/uk news/how hackers-can steal-your 8576657, accessed on 12 August 2022.

[36] Al-Garadi, M.A., Hussain, M.R., Khan, N., Murtaza, G., Nweke, H.F., Ali, I., Gani, A. (2019). Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges. IEEE Access, 7: 70701-70718. https://doi.org/10.1109/ACCESS.2019.2918354

[37] Kowalski, R.M., Giumetti, G.W., Schroeder, A.N., Lattanner, M.R. (2014). Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. Psychological Bulletin, 140(4): 1073.

[38] Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. Computers in Human Behavior, 23(4): 1777-1791. https://doi.org/10.1016/j.chb.2005.10.005

[39] Tokunaga, R.S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. Computers in Human Behavior, 26(3): 277-287. https://doi.org/10.1016/j.chb.2009.11.014

[40] Hinduja, S., Patchin, J.W. (2018). Cyberbullying Identification, Prvention, and Response. Cyberbullying, Research Center.

[41] Harms, C. (2007). Grooming: An operational definition and coding scheme. Sex Offender Law Report, 8(1): 1-6.

[42] Girouard, C. (2001). The National Center for missing and exploited children. US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

[43] Irshad, S., Soomro, T.R. (2018). Identity theft and social media. International Journal of Computer Science and Network Security, 18(1): 43-55.

[44] Wright, M., Kapadia, A., Kumar, M., Dhadphale, A. (2010). ReDS: Reputation for directory services in P2P systems. In Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research, pp. 1-4. https://doi.org/10.1145/1852666.1852747

[45] Ananthula, S., Abuzaghleh, O., Alla, N.B., Chaganti, S.B., Kaja, P.C., Mogilineedi, D. (2015). Measuring privacy in online social networks. International Journal of Security, Privacy and Trust Management, 4(2): 1-9. https://doi.org/10.5121/ijsptm.2015.4201

[46] Dey, N., Babo, R., Ashour, A.S., Bhatnagar, V., Bouhlel, M.S. (2018). Social Networks Science: Design, Implementation, Security, and Challenges. Springer International Publishing. https://doi.org/10.1007/978-3-319-90059-9

[47] Das, N., Borra, S., Dey, N., Borah, S. (2018). Social networking in web based movie recommendation system. In Social Networks Science: Design, Implementation, Security, and Challenges, 25-45. https://doi.org/10.1007/978-3-319-90059-9_2

[48] Potadar, D. (2021). https://github.com/dhavalpotdar/detecting-offensive-language-in-tweets, accessed on 12 August 2022.

[49] Shelke, N., Chaudhury, S., Chakrabarti, S., Bangare, S.L., Yogapriya, G., Pandey, P. (2022). An efficient way of text-based emotion analysis from social media using LRA-DNN. Neuroscience Informatics, 100048. https://doi.org/10.1016/j.neuri.2022.100048

[50] Gautam, A.K., Bansal, A. (2022). Performance analysis of supervised machine learning techniques for cyberstalking detection in social media. Journal of Theoretical and Applied Information Technology, 100(2).

[51] Akronix, (2019). https://github.com/Akronix/data-breaches-viz, accessed on 12 August 2022.

[52] Muneer, A., Fati, S.M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. Future Internet, 12(11): 187. https://doi.org/10.3390/fi12110187

[53] Gautam, A.K., Bansal, A. (2022). A review on cyberstalking detection using machine learning techniques: Current trends and future direction. International Journal of Engineering Trends and Technology, 70(3): 95-107. https://doi.org/10.14445/22315381/IJETT-V70I3P211