Neural Processing Letters (2021) 53:843–861
https://doi.org/10.1007/s11063-020-10416-3

Online Social Network Security: A Comparative Review Using Machine Learning and Deep Learning

Chanchal Kumar[1,2] · Taran Singh Bharati[1] · Shiv Prakash[3]

Accepted: 21 December 2020 / Published online: 5 January 2021

Abstract
In the present era, Online Social Networking has become an important phenomenon in human society. However, a large section of users are not aware of the security and privacy concerns involve in it. People have a tendency to publish sensitive and private information for example date of birth, mobile numbers, places checked-in, live locations, emotions, name of their spouse and other family members, etc. that may potentially prove disastrous. By monitoring their social network updates, the cyber attackers, first, collect the user's public information which is further used to acquire their confidential information like banking details, etc. and to launch security attacks e.g. fake identity attack. Such attacks or information leaks may gravely affect their life. In this technology-laden era, it is imperative for users must be well aware of the potential risks involved in online social networks. This paper comprehensively surveys the evolution of the online social networks, their associated risks and solutions. The various security models and the state of the art algorithms have been discussed along with a comparative meta-analysis using machine learning, deep learning, and statistical testing to recommend a better solution.

Keywords Online social networks · Attacks · Security · Malware analysis · Feature extraction · Big data · Machine learning · Deep learning

1 Introduction and Motivation

Humans are social creatures and sharing their emotions, desires, happiness, hobbies, and interests, is an important aspect of their socialisation. In twenty-first century people are living in a small world confined by a scarcity of time, and yet want to connect with other people of similar interests locally and globally. Online Social Networks (OSN) have emerged as platforms to connect people through various applications to form a social network that surpasses geographical and political boundaries (Fig. 1). OSN is a web site designed to

✉ Shiv Prakash
shivprakash@allduniv.ac.in

[1] Department of Computer Science, Jamia Millia Islamia, Jamia Nagar, New Delhi 110025, India

[2] Ciena India Pvt. Ltd, Gurugram, Haryana 122002, India

[3] Department of Electronics and Communication, University of Allahabad, Prayagraj 211002, India

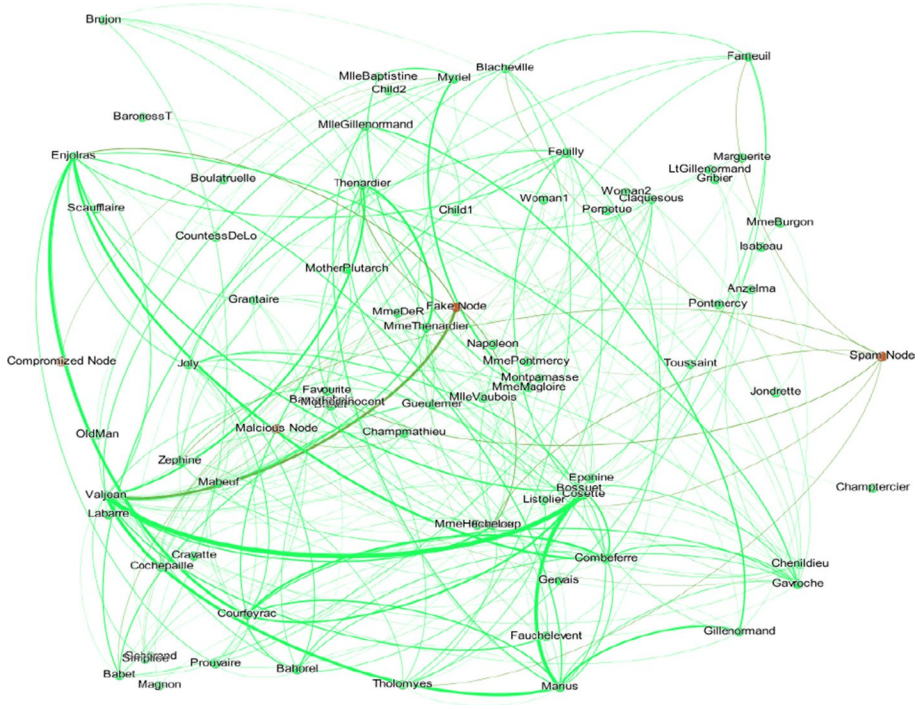

Fig. 1 Graphical representation of online social networking [3]

publish content freely on any subject to share with 'friends' and the public. A user in the OSN world is recognized through his profile known as his online social identity. The inception of Online social networking era is marked by the launch of SixDegree.com by Andrew Weinrich in 1997. From its humble beginnings, the online social network has developed from only a place to make friends to a big marketplace for business advertisements, sharing and promoting multimedia contents, and many more public relations services to unite the community. Presently all the top brands are highly active on social media platforms to promote their businesses and products. Statista in 2016, published a report showing the incredible increase in social network data being generated online. Besides the continuous advancement of Internet technologies [1], traditional social networks are also shifting to the online platform to provide efficient and effective services to their respective users. These technologies have become a gateway for the majority of people to connect with family, friends and colleagues across the globe. Today all the top brands are active on OSN, as per the report, globally more than 70% of businesses use OSN for marketing, but few are aware of the optimal ways to use OSN and of the tools keep themselves safe and secure in the virtual world. For example, in early 2020, there was a serious breach in the hotel Marriott's database which remained undiscovered for approximately a month. The month of April 2020 witnessed an accident of ransomware cyberattack and data breach. The Magellan healthcare was hit by a ransomware attack which was launched by installing malware in the system to steal the credentials of the employees. In mid-2020, hackers attacked the account of some high-profile individuals of the USA on Twitter and were able to reset their

account passwords. Most recently, when they work from home becomes the new normal and people had started using meeting apps for their virtual meeting, Zoom, which is one of the most frequently used video conferencing apps became a victim of privacy breach. The passwords of around five lakhs accounts belonging to various companies, institutions, schools and other organizations were put on sale [2]. In this virtual world where people, big institutions, government, and non-government organizations are connected virtually, issues of privacy, and data security have emerged as major concerns for everyone.

1.1 Chronology of OSN

The social network is an interdisciplinary word was in existence a long time before it was introduced in computer Science. In 1973, Douglas Brown and Dave Woolley developed a multi-user chat room application called Talkomatic. However, the Six Degree by Andrew Weinrich in early 1997 is considered as first officially OSN [4]. Since then a larger number of social networking site has been launched. The Table 1 depicts some of the popular OSN sites [5].

1.2 Type of OSN

Based on the prime service and interest of a user, OSN can be broadly classified in the following categories [6, 7] (Table 2).

1.3 Types of Malicious Profile

In [5, 8] author classified the user's profile based on the characteristics and activities which are as following:

- *Spam-profile* Spread voluntary/Unwanted messages over OSN.
- *Fake-profile* Many users create a bogus account to perform different malicious activities. The data collected by them is either copied information from the real user or the pseudo name are generated using automated programs which further used to create a fake account. Such Profile is also known as a sybil or cloned profile.
- *Compromised-profile* A profile hijacked from a legitimate user to spread various malware.

Figure 2 depicts that the Spam account first establishes connection/friendship with other accounts to win the trust of the user then it starts to send some unsolicited messages to other connected users. The figure also portrays how a malicious user steals some of the information from an authentic user and make a similar online profile (v5) to get connected with his/her known user (v4).

The malicious profile V5 sends a request to connect with V4, as V4 and V1 have so many common links, V5 has some information which is common to V4, therefore, V4 accepted the request. Once the request has been accepted, malicious user V5 starts collected the information like the visited place, likes, common friends, etc. Later, this information can be used in many ways to harm V4 or other similar users.

Table 1 Chronology of OSN

Year	Name of OSN site	Year	Name of OSN site
1997	SixDegree	2006	Twitter
	AsianAvenue		Tuenti
	CaringBridge		MyChurch
1998	Care2		Jaiku
	Xanga	2007	Bahu
	Open Diary		Tumblr
	Fotki		Fuelmyblog
1999	Advogato		Flixster
	Cyworld	2008	ResearchGate
	LiveJournal		Academia.edu
	BlackPlanet		MeinVZ
2000	Faceparty		Gays
	Habbo	2009	Skoob
	Trombi		DailyBooth
	DeviantArt		Foursquare
2001	Partyflock		Sina Weibo
	Kwick	2010	Jiepang
	Jappy		Diaspora
	CozyCot		Instagram
2002	Skyblog		Friendica
	Fotolog	2011	Gentlemint
	Friendster		Wellwer
	Reunion		Pinterest
2003	Xing		Google+
	MySpace	2012	Spot.IM
	Netlog		Sgrouples
	LinkedIn		Stage 32
2004	Facebook		Cucumbertown
	Orkut	2013	Smartican
	Flickr		Spring.me
	Dogster	2014	Ello
2005	Yahoo!360		Poolwo
	MocoSpace	2015	Blab
	Renren		Periscope
	Ning	2016	Mastodon
			Gab.ai

1.4 Type of OSN Security Risks and threats

The people used to post a lot of personal information which seems trivial to them. However, such information may become sensitive and severely harmful when it is exploited by an attacker with proper planning. The risk associated with OSN can further be classified as following [9]:

Table 2 Types of online social networks

Type	Focus	E.g.
Profile-based Social Networks	Revolve around personal profile which consists of personal information, likes, cultural and political interests, etc.	Bebo, Facebook and MySpace
Content-based Social Networks	Groups, community and comments are focused around content shared on the site	Flickr, YouTube.com, last.FM, Shelfari
White-label Social Networks	Some of the websites offer only the group and community building service to the users	Wikipedia, Wikileaks, PeopleAggregator, Ning
Micro-blogging or Presence Updates	A site providing short message service to update news and their personal views publicly	Twitter, Jaiku, NairaLand, Wayn

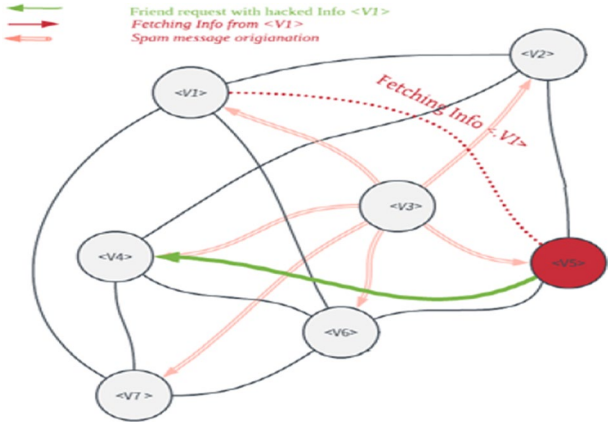

Fig. 2 Security in social networks [5]

1. *Risk related to the organizations* The organizations use a social network for personal or official purposes. Any severe attack on any such OSN may damage the entire network of the organization.
2. *Risks to the people* It may possible that the people using any network dealing with sensitive personal information may accidentally or intentionally expose their personal information on their social network.

The threats to OSN can be categorized as following [10]

1. *Classic threats* Threats that are not only specific to OSN but also exist in the Internet, e.g. Malware, Phishing attacks etc.
2. *Modern threats* Threats that exploit OSN infrastructure to breach security and privacy. E.g. Identity Clone Attacks, Clickjacking.
3. *Children specific threats* Threats that targets children over a social network e.g. cyberbullying, online predators.
4. *Combination threats* Threats which are launched by combining classic and modern threats.

2 Literature Review and Related Work

In the modern era, malicious activities are increasing massively in OSN. We have categories the literature review based on standard security requirements available i.e. confidentiality, integrity, and availability.

In [11] author discussed the phishing attacks which is an online identity stealing technique using societal and technical features to hijack user credentials e.g. user-names, passwords, etc. A phishing attack can be launched using various methods but the prevalent one takes phishing message reaches in the user mail-box fantasizing to be from a bank, directing the user to a web-page and asking him to enter his authentication, but the web-page is not one related with the bank. In this web-site phishing is deliberated, in which available solutions are based either on providing early caution of mistrustful activity and quick reply or on the usage of TLS (Transport Layer Security). In [12], the researcher presented phishing attacks, based on identity theft pointed to acquire sensitive information

e.g. authentication information, credit and debit card information, etc., by using masquerade attack as a trusted user in the OSN. Phishers applied various techniques, for example, e-mail spoofing to attempt to trick their victims. Though several anti-phishing techniques are available, yet phishers endure succeeding to trap victims. A novel technique that aims at classification a forged OSN submitted random Ids before the actual Ids in the login in OSN. A mechanism to analyse the replies from the server beside the proposals of each authentication information to decide if the OSN is unusual or phished. However, this technique is generalized and may work in each authentication process which is constructed on interchange of each credential, the existing prototypes are established for OSN HTTP Digest Authentications and to accept user-ID and password.

In [13] author presented system Lockr, a system to improve the privacy of content shared in centralized and decentralized systems. This method offers three substantial privacy advantages to OSN users. Firstly, it detached content to be shared from the functionality of OSNs to control its privacy of shared information. The system enabled the user to have control regarding access and storage of content y OSN provider or any 3rd party to accommodate user's privacy preferences. Any digitally signed OSN relationships required to access content can't be used for any inadvertent purpose. The importance of other's social content available to OSN providers reduced drastically using this feature. Finally, a social relationship key is used to encrypt a message which enables a common friend between two strangers to authenticate their relationship without exposing any information. In [14], the author discovered irresponsible activities that lead to the misuse of OSN. The paper emphasis on privacy paradox which states the conflicts between user actual behaviour and security awareness. In this paper, users' activities that lead to inadequate protection of circulated sensitive information were identified. Such sensitive information is enough to launch each type of phishing attack. In [15] author proposed an approach to use bogus bites to defend against phishing attacks instead of validating only phishing websites. Bogus Biter, A unique client-side anti-phishing tool, is developed. The Bogus Biter feeds a large number of fake credentials to a suspected website. The tool can conceal the real credentials of a user and equipped a legitimate website to detect stolen credentials rapidly. The tool was developed as an extension of a web browser. In [16] author referred to an Edge-Rank technique applied by Facebook where a score based on some selected features (likes, comments, reposts, remarks, etc.). The greater Edge-Rank scores the less possibility to be a spam user. However, there is a limitation of applied techniques that the spammer can boost their post score with the help of colluding networks or users. Currently, Twitters trends have become a popular tool to influence and build an opinion about a particular topic, product, and people. The manipulation in a trend may mislead the people. Therefore, it is necessary to discuss the security of the Twitter trend. In [17] the author discussed the manipulation of Twitter trends and emphasis on the investigation of the security of Twitter trends through the data analysis. The evidence of trend manipulation is collected by applying the influence model over hash-tag. The trends are studied topic wise based on their coverage, potential transmission, popularity, and reputation. The SVM classifier is used to know how precisely these factors could predict a trend. The author in [18] mentioned various challenges to identify influential bloggers and proposed a model to find out the most influential bloggers. They also found out the present activeness of these bloggers. In [19] the author described the information patterns and their privacy implications in OSN. The online behaviour of around four thousand students was taken into consideration. The study is carried about the uses of privacy settings and the amount of information disclosed by the user. Further in this, they emphasised promising attacks on some aspects of user's privacy with a minimum percentage alteration to the enormously permeable preferences of privacy.

In [20] the author focused on a classification technique based on supervised machine learning which focuses to detect malicious content spreading in OSN. For this purpose, multi-source features are used to identify the post which contains malicious Uniform Resource Locators (URLs) which further launch attacks e.g. phishing attack, spam, and scams. To address this issue, the author used the random forest (RF) classification technique without a feature selection method and any tuning. The recall value of the developed model is 89%. Further with the use of feature selection methods and tuning, the recall value increased to 92%. The malicious attackers are not only interested in information of a specific profile but also hacked OSN website. The author took up this issue in [21] and introduced new term antisocial networks. The Antisocial Networks are a distributed system of websites to exploit for various attacks. An adversary hacked visitor's session control, remote manipulation their browsers via a legitimate web-control function e.g. loading image, HTML tags, J-Script, and Java applets.

In [22] the author surveyed the approach and potential challenges to fight against spam on OSN. In this survey paper, three main anti-spam strategies i.e. demotion, detection and prevention. Detection based strategy focused on spam detection and eliminates its impact. Demotion based strategy tries to reduce the rank of spam while the prevention-based strategy attempts to minimize content contributing to spam by limiting user actions or interface change.

In [23] studies the concept of measurable trust management policies and protocols that minimize the chances of the discloser. Two users require a minimum total threshold point before granting access to a resource. The first user values certain credentials with point and grant access to another user only when the threshold value of points has been achieved while the second user values each credential with a privacy score which measure the degree of unwillingness to disclose that credential. The first user's valuation score of credentials and his threshold as well as the second user's privacy-valuation of credentials both are private. The user who wants to access resources tries to find out a subset of credentials that attain the threshold value required by another user. This paper demonstrated the protocol used for computation e.g. a subset of user's identifications without tightfitting any of the two users with private information. Besides this, a fingerprint method is used to permits the user an independent and easily recovery of the optimal solution for the knapsack problem. The fingerprint process is valuable afar the specific authorization deliberates and may be used in the private setting by using knapsack and dynamic-programming. In [24] author surveyed the recent threats to social networks and the attack-prone areas. The social networks are categorized into two sections user and social networking sites. Along with the countermeasures against the threats and proposed a security framework of social networks. In [25], the author analysed the footprints by collecting the data from ten popular social networking site. The data of approximately 14,000 active users were considered to gather information. After analysis of gathered information, it was found that a user active on one social network on average revealed 4.3 personal information fields while the user active on more than 8 social networks, this value increased up to 8.25 fields. Further, it was found that the attacker was able to reconstruct more than 40% of an individual footprint even with a single field. In [26], the author elaborated on security and privacy issues of OSN with some essential design clashes between old design objectives e.g. sus-ability and sociability. The core functionalities of OSNs were analysed to highlight security and privacy design challenges and opportunities to utilize OSN theory to reduce these design conflicts.

Most of the backend server used for data storage used the Linux operating system. Therefore, malware attack has started to target these systems. Most of the traditional malware detection methods developed so far may not be capable to defend against this

malware. In [27], the author introduced a new technique to recognize malicious executable linkable files. A system call tracer "Strace" is used dynamically to extract system calls. The various classification methods are applied to data sets using different features. The applied approach results in 97% classification accuracy. In [28] author proposed another group of attacks to breach user privacy. This novel class of attack exploits advertising systems having the capability to the target audience at the micro-level. The author targeted to study the advertisement system of Facebook, one of the world's most popular and large online social network systems. Several design choices of advertising systems have been identified and analysis which may result in information leakage through various advertising campaigns. Further, it has also been pointed out why the fixes provided by Facebook to address the mentioned vulnerabilities were insufficient to defend against.

Now a day a person has several accounts on the social network. Depending upon the type of service it is providing, the level of security and privacy varies from one online network to others. In [29], the author claimed that the network which has the weakest privacy rules in the OSN would determine how much personal information of any user is disclosed online. Using these measures, it is possible to estimate the user's social footprint vulnerability specifically towards security attacks e.g. physical ID and recovery of password. In this paper, the researcher demonstrated information leakage through data crawling over a public network and suggested some solutions for better privacy protection. In [30] the author presented a new model by applying a supervised machine learning algorithm. The data is collected through crawling from sina-weibo and then extreme learning technique applied using the user's behavioural characteristics to detect spam accounts. In the past decade, due to email spam detection and classification mechanism are in prime focus. In [31], the author summarized the main work done in the area of spam detection into two categories i.e. the Identity-based model and the content-based model. In the first model, the content is parsed according to spam potential patterns and keywords. In the identity-based model, email addresses are grouped into two lists i.e. whitelist and blacklist. The anti-spam mechanism worked as per these lists and decide which email address should be blocked and which are not. In [32], the author talked about content filtering in social interactive data. A Cloud-based Trust Awareness and Interaction Model (CTAIM), composed of Bayesian inference algorithm in Dirichlet distribution and Bayesian content filtering algorithm, developed to do content filtering of interactive data to find trust evaluation with high-efficiency, security, and neutrality. Based on interaction history and node behaviour, the proposed model is capable to provide 3rd party trustworthiness evaluation. In [33], the author took to Twitter and proposed a novel spam classification algorithm based on Bayesian to distinguish suspicious behaviour of the user with an F-measure value(precision result) of 89%. In place of examining each spam message individually, the author in [34] used a set of novel features to reconstruct spam messages and achieve a precision value over 80%. In [35], the author depicted about systematic look at existing privacy-preservation techniques and highlights the vulnerabilities of users. In [36] author emphasizes risks consideration by OSN companies to individual's information privacy while using the analytics process. OSN companies should develop accountable measures to attend its use. In [37] the author talked about the one-dimensional relationship between the personal needs of privacy and self-disclosure of information with each other or are independent. In [38], the author discussed Friend-in-the-middle attacks on social networking sites which can be used to harvest social information in an automated fashion.

A wide range of false data, particularly gossip data, has saturated relatively every side of informal communities. Consequently, programmed appraisal of information validity has gotten significant consideration. In [39] the author discussed the detection of rumours on

OSN. To distinguish rumours from normal messages, the author proposes a rumour detection method based on implicit features of user and posted contents. This paper addressed issues of rumour detection in microblogs. To identify a rumour three categories of features i.e. network-based, content-based and micro-blogs-specific memes are explored. With the help of these features, it is also possible to identify disinformers who endorse to originate and spread rumours. Most of the social-spam messages contain a URL to spam-content, malware, or pornographic websites [40]. Any activity results in irregular behaviour or response are considered as an anomaly which can be either static or dynamic, labelled or unlabelled. In [41], the author surveyed methods for detecting various types of anomalies. The Author in this research work demonstrated the two-stage anomaly detection process. In the first stage network features are selected and processed while the classification of observations from this feature space is done in the second stage. The anomalous behaviour if not rectify quickly may be a major contributor to network structure formation. For example, in an online auction system, fraudulent individuals may collaborate to boost their reputation. This paper surveyed existing computational techniques for detecting anomalies. In [42], the author proposed an evolutionary algorithm that used to form dynamic social networks for link prediction. In the above stated link prediction work is not using previous background knowledge so that we can apply available models to predict with better accuracy.

In [43] author, proposed a novel GA to detect community in OSN. This algorithm implements matrix encoding to enable traditional crossover between individuals. To enhance the diversity of initial individuals, initial populations are generated using similarity of nodes and retaining an acceptable level of accuracy. This improves the efficiency of the optimal solution search. The genetic operators such as crossover mutation and inversion are applied to achieve targeted solutions. In this paper, a highly parallelizable seed-based greedy algorithm is developed for the detection of various communities as possible from the weighted entity for the consistency network.

In [44], the author introduced an automated and behaviour-based malware analysis and labelling system, AMAL, to expose limitation of the existing system. AMAL is composed of two subsystems known as MaLabel and AutoMal. AutoMal is used to gather low granularity behavioural artifices and MaLabel used these artifices to generate features. In [45], the author focused on the profile replicating where the aggressors can match a user's present OSN and, so, trusted relations are made by using fake-profile. Separate OSN may filter spam they obtain, though they usually want huge resources (e.g., personnel) and incur a delay before to detect novel types of spam. A structure is applicable to detect spam applicable to each OSN. Other techniques (e.g. black-lists and message shingling) may be combined and centralized [46, 47].

In [48] author, surveyed the present security issues pertain to wide-spread OSN. Some issues related to privacy concern, viral-marketing, structural-attacks and malware-attacks are discussed with a focus on the privacy concerns. Various techniques Twitter data sentiment analysis and a comparative study of those techniques. Sentiment analysis of Twitter information is used to analyse the tweets, whether it is positive, negative or neutral [49]. They gave the meanings of various terms that are used instead of sentiment such as opinion, view, belief, etc. Their results showed that methods such as naive Bayes and SVM have the highest accuracy.

OSN platforms are most prevalent in the modern world because many people use several forms. In this platform, users allow connecting with their friends, classmates, colleagues, and family to share their personal views and information. Though, the issues associated to maintain the privacy, integrity, and security information available in various forms including multi-media, e.g. photos, videos, and audios, therefore, there

is a huge scope of the security breach. In [50], a multi-population cultural technique (MPCT) for community detection in OSN is proposed by the author. The communities can be identified by using available techniques which may categories into, loose and compact communities. Further, the Link prediction in OSN has several possible applications, for instance, recommended new items to users, friendship suggestions and discovering spurious connections [51]. The present the link prediction work is based on a meta-path-based technique by using, support vector machines (SVM) technique and K-nearest neighbour (KNN) techniques) with an average accuracy of 89%. In [51, 52], the author developed techniques to predict link on similarity-based techniques, maximum likely-hood techniques, and techniques based on probabilistic techniques. Usually, link-prediction techniques process a set of features for learning and prediction whether two users in data may be connected. In [53], the author addressed the influence of the maximization problem in social networks. The primary focus is to obtain a small subset of nodes maximizing influence. The problem in using greedy algorithms, Stochastic optimization algorithms such as simulated annealing in solving influence maximization problem has also been addressed. A novel GA is used to maximize influence by using through multi-population to maintain diversity. In various cases, modified GA is affected its structure which guarantees the diversity in the available solutions, the optimization process turns into more sophisticated. In the future emphasis may be on the change in the structure of the given population with attributes of the influence OSN. In [54], a hybrid technique with SVM and Whale Optimization Algorithm(WOA) is applied to detect spam profiles in the OSN platform. In this model, research gap is that during the spam feature selection, we can use the latest technique such as twin SVM etc. and we can also apply the efficient meta-heuristic and hyper-heuristic to increase the accuracy.

2.1 The Problem and Proposed Comparative Meta-Analysis

In order to achieve OSN security, the data is encrypted which results in the generation of voluminous data also referred to as big data. Online conversions may involve various multimedia consisting of texts, images, audio, and video. The processing of such voluminous data with traditional computing is not possible in a reasonable time with the desired accuracy, efficiency and reliability. Therefore, it may use High-performance computing (HPC) to process such voluminous data [55, 56]. The HPC may be used supercomputers and parallel processing algorithms to solve complex problems with huge data. In the literature, the researchers used different algorithms for this purpose. The quality of service (QoS) parameters such as i.e. Accuracy (A), Precision(P), Recall(R), and F Score(F) are considered to compare available algorithms. It is not possible to compare an algorithm with any other existing algorithms which involves more than one QoS parameter, therefore to address this problem, a novel framework is developed by using statistical testing(ST) such as *t* test, Chi square test, Wilcoxon etc. machine learning and deep learning techiniques [57].

A generic framework for comparative analysis of various techniques is shown in Fig. 3. The performance of these may vary depending upon the data sets being used, as well as on the suitability of technique to the data and the application domain under consideration (Fig. 3). Hence, determining which technique is most suitable for a specific

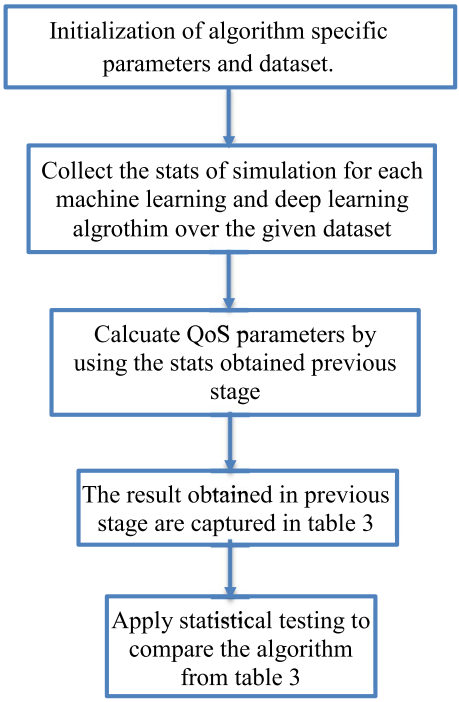

Fig. 3 Flow diagram of general framework

Table 3 Comparative result analysis of various techniques

Name of algorithm	Recall (%)	Precision (%)	Accuracy (%)	F Score (%)
Support vector machine (SVM)	95.00	93.12	96.90	94.05
k-nearest neighbour (KNN)	97.14	87.00	96.20	91.79
Artificial neural network (ANN)	96.92	96.02	96.83	96.47
Artificial immune system (AIS)	93.68	97.75	96.23	95.67
Rough sets (RS)	92.26	98.70	97.42	95.37
Relevance vector machine (RVM)	96.00	95.00	96.90	95.50
Random forest (RF)	90.00	97.80	97.00	93.74
Logistic regression (LR)	65.70	93.60	93.00	77.21
Naïve Bayes(NB)	83.00	86.00	87.00	84.47

application domain and its related data sets would be a key advantage. The summary of the results obtained is given in Table 3.

2.2 Parameters Used in Comparative Result Analysis

True = T; False = F; Positive = P; Negative = N;

Accuracy (A): Accuracy is measured how close a measured value (M) and actual (true (T)) value.

$$A = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{1}$$

Precision (P): Precision (P) is measured how close the measured (M) values are to each other.

$$P = \frac{TP}{TP + FP} \tag{2}$$

The Precision signifies how precise/accurate your model. Out of those all predicted positives, how many of them are actual positives. When the cost of False Positive is high, it good to determine Precision value. For instance, in reference to email spam detection, a false positive signifies a non-spam email (actual negative) has been recognized as spam (predicted spam). The email user might lose important emails if the precision is not high for the spam detection model.

Recall (R): Recall (R) is the portion of the relevant documents which are successfully recovered.

$$R = \left(\frac{TP}{TP + FN}\right) \tag{3}$$

F Score (F): It is defined as a harmonic mean of precision and recall.

$$F = 2 * P * \frac{R}{(P + R)} \tag{4}$$

The time complexities are constructed on user experience and the skills with implementation number of instances of data are n, having m features such that n is greater than m, e = no of iterations, k = number of neurons/number of clusters, M = number of trees. According to the thumb rule, the time complexities O(n) and O(nlog n) techniques are assumed as linear time and used to online methods. Further $O(n^2)$ is assumed as tolerable to most of the available real-time common problems. Therefore $O(n^3)$ and higher are assumed to be far slower techniques and used for offline methods. The time complexities of techniques are Naïve Bayes = O(mn), ANN = O(emnk), KNN = O(nlogk), Random forest (RF) = O(Mmnlogn), SVM = $O(n^2)$. Therefore, due to different time complexities, requirements, and varieties, we used the statistical techniques to know which technique is better in what attribute and feature. The big data available in various security techniques in the OSN required high-performance computing techniques.

2.3 Statistical Testing to Compare Existing Algorithms

A set of ST tests like *t*-test, *z* test, *F* test; Chi square test, etc. are used to test between the given sample and the hypothetical or expected samples to obtain the significant levels of consequences. As the statistical testing considers all parameters used in any technique therefore it is used to get deep knowledge about the performance of available techniques [57]. These tests are used in a pair to compare the performances of the two techniques.

In this paper, different techniques are compared and analysed using the null hypothesis i.e.$H_0 : \theta_D = 0$ where the difference between hypnosis and confidence interval is zero for different samples and the average of the various scores is zero. If T is less than or equal to the value of the distribution of a statistical test, the null hypothesis is rejected. The *p*-value

Table 4 Result of Wilcoxon test for KNN Vs ANN

+tive rank	−tive rank	Ties	R^+	R^-	S	*p*
19	7	4	274.50	76.50	− 2.516	0.012
28	2	0	432.00	33.00	− 4.103	0.000
29	1	0	460.00	5.00	− 4.679	0.000
14	9	7	177.00	99.00	− 1.187	0.235
25	5	0	401.00	64.00	− 3.466	0.001
28	2	0	449.00	16.00	− 4.453	0.000
10	10	10	106.00	104.00	− 0.037	0.970
21	8	1	333.00	102.00	− 2.498	0.012

Table 5 Result of Wilcoxon test for SVM Vs KNN

+tive-rank	−tive rank	Ties	R^+	R^-	S	*P*
27	2	1	417.00	18.00	− 4.315	0.000
12	12	6	155.00	145.00	− 0.143	0.886
16	12	2	216.50	189.50	− 0.307	0.758
20	9	1	308.50	126.50	− 1.968	0.049
27	3	0	426.00	39.00	− 3.980	0.000
29	1	0	462.00	3.00	− 4.730	0.000
9	8	13	82.50	70.50	− 0.285	0.776
24	5	1	382.00	53.00	− 3.557	0.000

related to a comparison is obtained using the normal approximation for the statistical test. The level of significance α is selected as 0.05 which states that if the *p*-value is greater than α, then there is no major difference between these techniques. In this paper, various tests are applied; however, it found that the results obtained through Wilcoxon tests are most effective to perform a meta-analysis.

2.4 Wilcoxon test

It is two-tail tests where R^+ represents the first tail while R^- represents the second tail. If the value of R^+ is greater it means the first tail is performing better while the greater value of R^- represents that the second tail is better. Further, if both R^+ and R^- value are the same then both techniques are equally good. The results of statical testing are given in Tables 4, 5, 6 and 7.

In this case, the test data set is taken in multiple of 30 e.g. 300/3000 which can be further taken up to 300,000 depending upon the available computing facility. KNN technique is the first tail and the second tail is ANN. In Wilcoxon, test R^+ represents the first tail while R- represent the second tail. Here R Positive is greater therefore the first tail i.e. KNN is performing better. The corresponding results are obtained by implementing the proposed framework on Dell WorkStation, Intel Xeon E5–262 v4 processor, 128 GB RAM, and Windows 10 operating system.

Table 6 Result of Wilcoxon test for RF Vs SVM

+tive-rank	−tive rank	Ties	R^+	R^-	S	p
27	2	1	417.00	18.00	− 4.315	0.000
12	12	6	155.00	145.00	− 0.143	0.886
16	12	2	216.50	189.50	− 0.307	0.758
20	9	1	308.50	126.50	− 1.968	0.049
27	2	1	417.00	18.00	− 4.315	0.000
14	15	1	173.00	262.00	− 0.963	0.336
9	20	1	141.00	294.00	− 1.656	0.098
12	12	6	155.00	145.00	− 0.143	0.886

Table 7 Result of Wilcoxon test for RF Vs RVM

+tive-rank	−tive rank	Ties	R^+	R^-	S	*p*
23	5	2	354.00	52.00	− 3.439	0.001
13	6	11	106.50	83.50	− 0.463	0.643
17	13	5	254.50	210.50	− 0.453	0.651
10	14	6	110.50	189.50	− 1.129	0.259
14	15	1	173.00	262.00	− 0.963	0.336
9	20	1	141.00	294.00	− 1.656	0.098
12	12	6	155.00	145.00	− 0.143	0.886
24	5	1	382.00	53.00	− 3.557	0.000

In this case, the test data set is taken in multiple of 30 e.g. 300/3000 which can be further taken up to 300,000 depending upon the available computing facility. SVM technique is the first tail and the second tail is KNN. In Wilcoxon, test R^+ represents the first tail while R^- represents the second tail. Here R Positive is greater therefore first tail i.e. SVM is performing better.

Finally, we determined the state of art results and compare our results with the state of results. In this case, the test data set is taken in multiple of 30 e.g. 300/3000 which can be further taken up to 300,000 depending upon the available computing facility. RF technique is the first tail, and the second tail is SVM. In Wilcoxon, test R^+ represents the first tail while R^- represents the second tail. Here R Positive is greater therefore first tail i.e. RF is performing better.

In this case, the test data set is taken in multiple of 30 e.g. 300/3000 which can be further taken up to 300,000 depending upon the available computing facility. RF technique is the first tail, and the second tail is RVM. In Wilcoxon, test R^+ represents the first tail while R^- represents the second tail. Here R Positive is greater therefore first tail i.e. RF is performing better.

Finally, after comparing the result data available in the Tables 4, 5, 6 and 7, we can say that RF, is the best technique for the spam filtering in the OSN. Besides this comparative study is also given in Table 8.

Table 8 Summary of compared work

Approach	Advantages	Future directions
Michael Fire [10]	Presented of the different security and privacy risks and solutions with special attention to children specific threats	Discovering novel types of security and privacy threats are suggested for future work
Prateek Joshi [58]	Mathematical and computational models for security and privacy are presented	Integration techniques with metrics will be considered in future work
Gail-Joon Ahn [59]	Various studies and recent news reports have highlighted	Implementation of different studies may be part of future work
Ferrag [60]	A survey of deep learning approaches for cyber security intrusion detection is presented	–
Framework of comparative analysis is proposed	The various security models and the state of the art algorithms have been discussed along with a comparative meta-analysis using machine learning, deep learning, and statistical testing	The optimization of available results can be combined with the available evolutionary and other approximation algorithms to analyse the security system in the available OSN

3 Conclusion and Future Directions

In the era of the cyber world, it is important to understand the criticality of the data shared over any social media platform and how to keep it safe and secure. In the literature, many privacy policies and privacy settings are provided by service providers as well as developed by researchers. The first step to achieve the desired security and privacy is to understand existing models. Along with the development of new models and techniques, there is a need to compare the existing models to adopt more robust and secure. In this paper, the framework is proposed and applied for a comparative analysis of available classification models in terms of QoS parameters and complexity. After meta-analysis, it is found the Random Forest with time complexity O (M*m*n*logn), is the best technique for the classification in the OSN. In the future, the proposed framework can be applied to compare some other machine learning and deep learning algorithms for real world problems. Further, the optimization of the implemented framework, the evolutionary, and other approximation algorithms can be used and scaled-up.

References

1. Rathore S, Sharma PK, Loia V, Jeong Y-S, Park JH (2017) Social network security: issues, challenges, threats, and solutions. Inf Sci (Ny) 421:43–69
2. Dutta P (2020) 5 Biggest Data breaches of 2020 (So Far). Security Boulevard, 2020 [Online]. Available: https://securityboulevard.com/2020/08/5-biggest-data-breaches-of-2020-so-far/
3. Hernández-García Á (2014) Using Gephi to visualize online course participation: a Social Learning Analytics approach. Ital J Educ Technol 22(3):148–156
4. Shu W, Chuang YH (2011) The perceived benefits of six-degree-separation social networks. Internet Res 21(1):26–45
5. Adewole KS, Anuar NB, Kamsin A, Varathan KD, Razak SA (2017) Malicious accounts: dark of the social networks. J Netw Comput Appl 79:41–67
6. Abdulhamid SM, Ahmad S, Waziri VO, Jibril FN (2011) Privacy and national security issues in social networks: the challenges. Int J Comput Internet Manag 19(3):14–20
7. Beye M, Jeckmans A, Erkin Z, Hartel P, Lagendijk R, Tang Q (2010) Literature overview-privacy in online social networks. Enschede
8. Stallings W (2017) Cryptography and network security: principles and practice. Pearson, Upper Saddle River
9. Salama M, Panda M, Elbarawy Y, Hassanien AE, Abraham A (2012) Computational social networks: security and privacy. In: Computational social networks. Springer, pp 3–21
10. Fire M, Goldschmidt R, Elovici Y (2014) Online social networks: threats and solutions. IEEE Commun Surv Tutor 16(4):2019–2036
11. Badra M, El-Sawda S, Hajjeh I (2007) Phishing attacks and solutions. In: Proceedings of the 3rd international conference on Mobile multimedia communications, p 42
12. Joshi Y, Saklikar S, Das D, Saha S (2008) PhishGuard: a browser plug-in for protection from phishing. In: 2nd international conference on internet multimedia services architecture and applications, 2008. IMSAA 2008. pp 1–6
13. Tootoonchian A, Saroiu S, Ganjali Y, Wolman A (2009) Lockr: better privacy for social networks. In: Proceedings of the 5th international conference on emerging networking experiments and technologies, pp 169–180
14. Nagy J, Pecho P (2009) Social networks security. In: Third international conference on emerging security information, systems and technologies, SECURWARE'09. pp 321–325
15. Yue C, Wang H (2010) BogusBiter: a transparent protection against phishing attacks. ACM Trans Internet Technol 10(2):6
16. Zheng X, Zhang X, Yu Y, Kechadi T, Rong C (2016) ELM-based spammer detection in social networks. J Supercomput 72(8):2991–3005
17. Zhang Y, Ruan X, Wang H, Wang H, He S (2017) Twitter trends manipulation: a first look inside the security of twitter trending. IEEE Trans Inf Forensics Secur 12(1):144–156

18. Agarwal N, Liu H, Tang L, Yu PS (2008) Identifying the influential bloggers in a community. In: Proceedings of the 2008 international conference on web search and data mining. pp 207–218
19. Gross R, Acquisti A (2005) Information revelation and privacy in online social networks. In: Proceedings of the 2005 ACM workshop on privacy in the electronic society, pp 71–80
20. Al-Janabi M, de Quincey E, Andras P (2017) Using supervised machine learning algorithms to detect suspicious URLs in online social networks. In: Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017, pp 1104–1111
21. Makridakis A, Athanasopoulos E, Antonatos S, Antoniades D, Ioannidis S, Markatos E (2010) Understanding the behavior of malicious applications in social networks. IEEE Netw 24(5):14–19
22. Heymann P, Koutrika G, Garcia-Molina H (2007) Fighting spam on social web sites: a survey of approaches and future challenges. IEEE Internet Comput 11(6):36–45
23. Yao D, Frikken KB, Atallah MJ, Tamassia R (2008) Private information: to reveal or not to reveal. ACM Trans Inf Syst Secur 12(1):6
24. Luo W, Liu J, Liu J, Fan C (2009) An analysis of security in social networks. In: Eighth IEEE international conference on dependable, autonomic and secure computing, DASC'09. pp 648–651
25. Irani D, Webb S, Li K, Pu C (2009) Large online social footprints—an emerging threat. In: 2009 International conference on computational science and engineering, pp 271–276
26. Zhang C, Sun J, Zhu X, Fang Y (2010) Privacy and security for online social networks: challenges and opportunities. IEEE Netw 24(4):13–18
27. Meraji S, Tropper C (2010) A machine learning approach for linux malware detection. In: Proceedings - international conference on parallel processing, pp 545–554
28. Korolova A (2010) Privacy violations using microtargeted ads: a case study. In: Proceedings - IEEE international conference on data mining, ICDM, pp 474–482
29. Irani D, Webb S, Li K, Pu C (2011) Modeling unintended personal-information leakage from multiple online social networks. IEEE Internet Comput 15(3):13–19
30. Liu C, Wang G (2017) Analysis and detection of spam accounts in social networks. In: 2016 2nd IEEE International Conference on Computer and Communications, ICCC. pp 2526–2530
31. Zheng X, Zeng Z, Chen Z, Yu Y, Rong C (2015) Detecting spammers on social networks. Neurocomputing 159(1):27–34
32. Xu L, Zheng X, Rong C (2013) Trust evaluation based content filtering in social interactive data. In: Proceedings - 2013 international conference on cloud computing and big data, CLOUDCOM-ASIA, pp 538–542
33. Wang AH (2011) Don't follow me - Spam detection in twitter," pp 142–151
34. Gao H, Chen Y, Lee K, Palsetia D, Choudhary A (2012) Towards online spam filtering in social networks. In: NDSS
35. Li N, Zhang N, Das SK (2012) Online social networks. In: Handbook on securing cyber-physical critical infrastructure found. Challenges, p 431
36. Schwartz PM, Solove DJ (2011) The PII problem: privacy and a new concept of personally identifiable information. NYUL Rev 86:1814
37. Shin S, Ko Y, Jang J (2011) The conflict between privacy and self-disclosure in social networking services. In: 2011 IEEE 3rd international conference on communication software and networks (ICCSN), pp 490–493
38. Huber M, Mulazzani M, Weippl E, Kitzler G, Goluch S (2011) Friend-in-the-middle attacks: exploiting social networking sites for spam. IEEE Internet Comput 15(3):28–34
39. Qazvinian V, Rosengren E, Radev DR, Mei Q (2011) Rumor has it: Identifying misinformation in microblogs. In: Proceedings of the conference on empirical methods in natural language processing, pp 1589–1599
40. Nguyen H (2013) State of social media spam. Publ. NexGate, USA, from websites http//nexgate.com/wpcontent/uploads/2013/09/Nexgate-2013-State-of-Social-Media-Spam-Research-Report.pdf
41. Savage D, Zhang X, Yu X, Chou P, Wang Q (2014) Anomaly detection in online social networks. Soc Netw 39:62–70
42. Bliss CA, Frank MR, Danforth CM, Dodds PS (2014) An evolutionary algorithm approach to link prediction in dynamic social networks. J Comput Sci 5(5):750–764
43. Pizzuti C (2014) GA-Net : a genetic algorithm for community detection in social networks GA-Net : a genetic algorithm for community detection in social networks. January 2008
44. Mohaisen A, Alrawi O, Mohaisen M (2015) AMAL: high-fidelity, behavior-based automated malware analysis and classification. Comput Secur 52:251–266
45. Ebenazer MP, Sumathi P (2015) An overview of identity deception approaches and its effects. Int J Comput Trends Technol 25(3):123–126

46. Multani HS, Sinh-Marod A, Pillai V, Gaware V (2015) Spam detection in social media networks: a data mining approach. Int J Comput Appl 115(9):9–12
47. Jeong S et al (2017) Social network security: issues, challenges, threats, and solutions. Inf Sci (Ny) 421(3):43–69
48. Saikumar T, Sriramya P (2016) Security issues in social networks. Int J Pharm Technol 8(4):20835–20841
49. Kharde V, Sonawane P et al (2016) Sentiment analysis of twitter data: a survey of techniques. arXiv: 1601.06971
50. Zadeh PM, Kobti Z (2015) A multi-population cultural algorithm for community detection in social networks. Procedia Comput Sci 52:342–349
51. Jalili M, Orouskhani Y, Asgari M, Alipourfard N, Perc M (2017) Link prediction in multiplex online social networks subject category : subject areas
52. Lü L, Zhou T (2011) Link prediction in complex networks: a survey. Phys A Stat Mech Appl 390(6):1150–1170
53. Zhang K, Du H, Feldman MW (2017) Maximizing influence in a social network: improved results using a genetic algorithm. Phys A Stat Mech Appl 478:20–30
54. Ala'M A-Z, Faris H, Hassonah MA et al (2018) Evolving support vector machines using whale optimization algorithm for spam profiles detection on online social networks in different lingual contexts. Knowl Based Syst 153:91–104
55. Mutlag AA, Abd-Ghani MK, Arunkumar N, Mohammed MA, Mohd O (2019) Enabling technologies for fog computing in healthcare IoT systems. Future Gener Comput Syst 90:62–78
56. Zhang Z, Choo K-KR, Gupta BB (2018) The convergence of new computing paradigms and big data analytics methodologies for online social networks. J Comput Sci 26:453–455
57. Ali M, Siarry P, Pant M (2011) An efficient differential evolution based algorithm for solving multi-objective optimization problems. Eur J Oper Res 217:404–416
58. Prateek K, Jay JC–C (2011) Security and privacy in online social networks : a survey. In: 2011 IEEE International Conference on Multimedia and Expo (ICME). Prateek Joshi, C–C Jay Kuo University of Southern California, Los Angeles, California, USA, pp 1–6
59. Altshuler Y, Elovici Y, Cremers AB, Aharony N, Pentland A (2013) Security and privacy in social networks. Secur Priv Soc Netw 15:1–253
60. Ferrag MA, Maglaras L, Moschoyiannis S, Janicke H (2020) Deep learning for cyber security intrusion detection: approaches, datasets, and comparative study. J Inf Secur Appl 50:102419