International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 03 | Mar-2018 www.irjet.net p-ISSN: 2395-0072

Competitive Analysis Of Attacks On Social Media

Narinderpal kaur[1], Rasleen Kaur[2]

[1]*Student, Dept of Computer Science Engineering, GIMET Amritsar, Punjab, India*
[2]*Assistant Professor, Dept of Computer Science Engineering, GIMET Amritsar, Punjab ,India*

Abstract - *Today social media play very important role in the communication process. With the help of social media number of tasks is being accomplished. Normally social media is used to sell products, Career Consultancy, Communication, sharing resources etc. Along with the advantages there are number of disadvantages of the social media also. Social media uses number of mechanisms to create users. And increase their database. But they do not ensure validity of information provided bsy the user. This will cause the deception. Deception will cause number of problems. There are number of types of deceptions which exist over the internet. Deception model is prepared in order to analyze these problems. Some of the deceptions are difficult to detect than the others. Some of the challenges which social media must address are considered in this paper.*

***Key Words*: Deception, social media, Internet**

1. INTRODUCTION

With the growth of the internet social media growth is increased.[7] The social media is used for wide variety of purposes. Social media is used to share user generated contents to large number of other users. With that the number of services provided by the social media also increases. With the advent of technology the deception also has been increased. Deception is caused due to falsifying information provided by the user.[3] Social media provide new environment for the deceivers to perform illegal tasks over the internet. The main cause of deception is that it is very easy to create account over the social media like Facebook, Twitter etc. No verification of records is done in case of the social media. They are just consider the increase of database and do not consider deception.In this paper we consider or attack deception as the deliberate attempt to provide falsifying information to conduct harm over the network.[31] The problem is extravagated since receiver does not know about the deception. Because of which receiver privacy will be on the stakes. The private information of the receiver will be determined by the deceiver. These false beliefs are transferred through verbal and non verbal communications.[18] Clone attacks are common source of deception over the social media. Rest of the paper is focused on determining clone attack detection strategies which result in redundant information or users over the social media. Today social media assume critical part in the correspondence procedure. With the assistance of social media number of errands is being refined. Ordinarily social media is utilized to offer items, Career Consultancy, Communication, sharing assets and so on. Alongside the points of interest there are number of burdens of the social media too. Social media utilizes number of components to make clients. What's more, increment their database. In any case, they don't guarantee legitimacy of data gave by the client. This will cause the double dealing. Misdirection will cause number of issues. There are number of kinds of double dealings which exist over the web. Double dealing model is set up keeping in mind the end goal to examine these issues. A portion of the double dealings are hard to distinguish than the others. A portion of the difficulties which social media must address are considered in this paper.

1.1 Machine Learning

To make the proposed counterfeit record identification framework adaptable; we composed and executed a commonsense machine learning pipeline including a grouping of information pre-preparing; highlight extraction; forecast and approval stages.[5] Machine learning calculations that have assumed a noteworthy part in social media examination incorporate Decision tree learning , Naïve Bayes, Nearest Neighbor classifier, Maximum Entropy strategy , Support vector machine(SVM) ,Dynamic Language Model classifier , direct relapse and calculated relapse , Simple calculated classifier , Bayes Net and Multilayer Perceptron.[18]The pipeline comprises of three noteworthy segments; which we depict beneath and show in Figure.

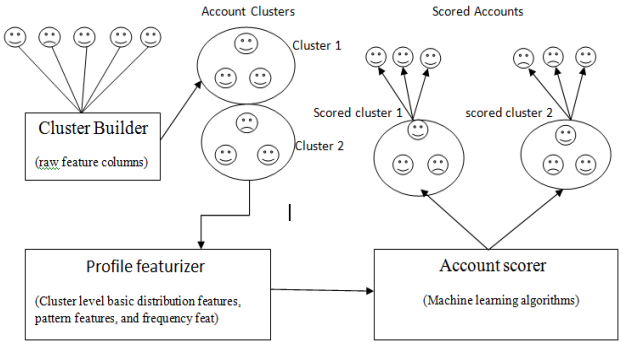

Fig 1.1 Machine Learing

The Cluster Builder, as its name suggests, takes the crude rundown of records and constructs groups of records alongside their crude highlights.[15]The module takes client specialized parameters for (1) least and greatest group estimate;(2) time traverse of records enlisted (e.g. most recent 24 hours, a week ago), and (3) grouping criteria. The grouping criteria can be as basic as gathering all records that

offer a typical singe act eristic, for example, IP address, or a more unpredictable grouping calculation, for example, k - implies.[8] Once the underlying groups are manufactured, client denied criteria can be added to later out a portion of the groups that are not prone to be suspicious or may present high false positives. For instance, one may wish to later out records enrolled from the OSN's corporate IP space, as these are probably going to be test accounts that ought not to be confined. The Cluster Builder takes crude part genius le tables as info and yields a table of records with highlights that are required for include designing, for example, part's name, organization, and instruction. Each line of the table speaks to one record and contains a "cluster identifier" novel to that record's group.[28] This table is utilized as contribution to the Problem Featurizer. In the preparation stage the Cluster Builder should likewise utilize account-level marks to name each group as genuine or phony. While most groups have either all records or no records named as phony, there will as a rule be a couple of groups with a few records in each gathering. Along these lines to process group labels, we pick an edge x to such an extent that the groups with less than x percent counterfeit records are named genuine and those with more prominent than x percent counterfeit are marked phony.[3] The operation optimal decision of x relies upon exactness/review tradeoffs (i.e., higher estimations of x increment accuracy to the detriment of review).[27]

1.2 Profile Featurizer

The Profile Featurizer is the key segment of the pipeline. Its motivation is to change over the crude information for each cluster (i.e. the information for the greater part of the individual records in the cluster) into a solitary numerical vector speaking to the cluster that can be utilized as a part of a machine learning calculation.[16] It is executed as an arrangement of capacities intended to catch however much data as could reasonably be expected from the crude highlights with a specific end goal to segregate clusters of phony records from clusters of honest to goodness accounts.[4]The separated highlights can be comprehensively gathered into three classes; which we depict at abnormal state here; additionally subtle elements can be found in Section.

Essential dispersion highlights. For each cluster; we take essential factual measures of every section (e.g. organization name).[24]Cases incorporate mean or quartiles for numerical highlights; or number of extraordinary esteems for content highlights. 2. Example highlights. We have outlined \pattern en-coding calculations" that guide client created content to a littler straight out space. We at that point take essential dispersion includes over these straight out factors.[22] These highlights are intended to identify malevolent clients (particularly bots) that are following an example in their air conditioning check information exchanges. 3. Recurrence highlights. For each component, we register the recurrence of that incentive over the whole record database. We at that point figure fundamental dissemination includes over these frequencies.[10]When all is said in done we expect clusters of genuine records to have some high-recurrence information and some low-recurrence information, while bots or malignant clients will indicate less change in their information frequencies; e.g., utilizing just normal or just uncommon names.[13]

1.3 Account Scorer

The Account Scorer's capacity is to prepare the models and assess them on already inconspicuous information. The Account Scorer takes information of Profile Featurizer; i.e., one numerical factor for each cluster is build.[20] The extraordinary learning calculation utilized is client configurable; in our investigations we think about strategic relapse, arbitrary woods, and bolster vector machines. In \training mode," the Account Scorer is given a named set of preparing information and yields a model portrayal and assessment measurements that can be utilized to think about various models.[9] In \evaluation mode," the Ac-tally Scorer is given a model depiction and an info vector of cluster highlights and yields a score for that cluster showing the probability of that cluster being made out of phony accounts.[6] In light of the cluster's score, accounts in that cluster can be chosen for any of three activities: programmed confinement (if the likelihood of being phony is high), manual survey (if the outcomes are uncertain), or no activity (if the likelihood of being phony is low).[19] The correct limits for choosing between the three activities are designed to limit false positives and give human analysts a blend of good and awful accounts. The physically named accounts can later be utilized as preparing information in advance emphases of the mode.[23]

2. BACKGROUND ANALYSIS

Techniques are devised to check the falsifying information provided by the user in order to achieve desired goals. These techniques are discussed in detail in this section.

2.1 Centralized Detection Technique

Grewal & Scholar 2015 Central authority is established in detection and prevention of clone attack in social media. Mainly this application is deployed in sensor networks but due to heavy traffic associated with the social media need for centralized detection comes into place. Under centralized detection following techniques appear.[14] Khabbazian et al. n.d.In this approach every node in the network send the information towards the neighbor and neighbor in turns sends the information towards the base station. The base station scans all the relevant information from the received report and if conflicting position information is spotted then message is conveyed to all the nodes in the network.[17] Solanki 2016 Another method of detection is the set operations. this mechanism reduces the number of transmitted packets hence overhead is considerably reduced by the use of set operations. In this mechanism duplicity

from the subset is detected and removed. This technique is cost effective but also disallows some of the critical packet transmission.[25] Wang & Zhang 2007 Cluster based approach is based on attribute similarity. The data being transmitted is analysed for similarity based on properties they have. In case of similar attributes disclosure, the information is packed in a common group known as cluster. Cluster contains information of similar sort hence these are homogeneous clusters. Threshold value upon the number of attributes similarity is maintained. In case similarity index exceeds threshold value, clone attack is detected. Heterogeneous cluster is not worked upon in existing literatures as yet.[29] Ren et al. n.d. This is the process in which key generated through cryptography process is analyzed for redundancy. The data transmitted in such fashion is secured and difficult to analyze. In order to detect replicated keys, extra storage is required at data centers. The upcoming key is compared against the incoming keys. The incoming keys if similar, replication is detected so does clone attack.[22]

2.2 Distributed Detection Techniques

These are the techniques which do not rely on the centralized authority to check for the abnormality. Watcher nosed are established in order to check for the anomalies. Under distributed techniques following mechanisms appear Devi & Poovammal 2016 Content based filtering mechanism is used to detect the abnormal material among the transmitted contents. The social media is prone to large number of users having large number of data associated with them. Content filtering mechanism maintains a word count register, containing the total number of words to be transmitted. After transferring the word, word count register is decremented by one. After word count register reaches 0, words to be transmitted still pending is analyzed. In case any word is still left, that indicates malicious entry. Some work towards saving time is still required to be accomplished.[11]Dave et al. n.d. Attributes are the properties associated with the content being transmitted. Attributes of use must be transmitted with integrity check. Primary key is implied over the attributes being transmitted. The attributes values cannot be redundant and also it cannot be null. Problem with the attribute based approach is attribute similarity is checked but content is not purified. [11]Mohammed et al. 2014 The attention on this paper is to assemble an Android stage based portable application for the medicinal services area, which utilizes the possibility of Internet of Things (IoT) and distributed computing. We have constructed an application called 'ECG Android App' which gives the end client perception of their Electro Cardiogram (ECG) waves and information logging usefulness out of sight. The logged information can be transferred to the client's private incorporated cloud or a particular restorative cloud, which keeps a record of all the observed information and can be recovered for investigation by the therapeutic staff. Despite the fact that building a restorative application utilizing IoT and cloud procedures isn't absolutely new, there is an absence of observational investigations in building such a framework. This paper surveys the major ideas of IoT. Further, the paper displays a foundation for the medicinal services area, which comprises of different advances: IOIO microcontroller, flag preparing, correspondence conventions, secure and proficient instruments for expansive record exchange, information base administration framework, and the concentrated cloud. The paper accentuates on the framework and programming engineering and outline which is basic to general IoT and cloud based restorative applications. The framework exhibited in the paper can likewise be connected to other medicinal services spaces. It finishes up with suggestions and extensibilities found for the arrangement in the human services space.[21] Tekieh & Raahemi 2015 In this review, we gather the related data that show the significance of information mining in social insurance. As the measure of gathered wellbeing information is expanding fundamentally consistently, it is trusted that a solid investigation device that is equipped for taking care of and dissecting extensive wellbeing information is basic. Breaking down the wellbeing datasets assembled by electronic wellbeing record (EHR) frameworks, protection claims, wellbeing studies, and different sources, utilizing information mining methods is exceptionally perplexing and is looked with certain difficulties, including information quality and security issues. In any case, the uses of information mining in social insurance, focal points of information mining procedures over conventional strategies, uncommon qualities of wellbeing information, and new wellbeing condition puzzles have made information digging extremely fundamental for wellbeing information examination.[26] Kiruthiga 2014 Social Networks (SN) are prominent among the individuals to associate with their companions through the web. Clients investing their energy in prevalent social systems administration destinations like facebook, Myspace and twitter to share the individual data. Cloning assault is one of the slippery assaults in facebook. Generally aggressors stole the pictures and individual data about a man and make the phony profile pages. Once the profile kicks cloned they off to send a companion ask for utilizing the cloned profile. In the event that if the genuine clients account gets blocked, they used to send another companion demand to their companions. In the meantime cloned one likewise sending the demand to the individual. Around then it was difficult to recognize the genuine one for clients. In the proposed framework the clone assault is distinguished in light of client activity era and clients click example to discover the comparability between the cloned profile and genuine one in facebook. Utilizing Cosine closeness and Jaccard file the execution of the similitude between the clients is made strides.[18].

Both content based and attribute based approaches commonly used with the applications of recommender system.

3. COMPARISON OF VARIOUS TECHNQIUES FOR CLONE ATTACK DETECTION

The comparison table for detection of clone attack is given as under

Authors and Year	Techniques	Attack Detected	Merit and Demerits
(Tsikerdekis & Zeadally 2014)[28]	Nonverbal Behavior	Multiple Identities Clone attack Detection	Non verbal behavior techniques is implied which gives result faster but it may not be accurate in all situations
(Egele et al. 2015)[12]	Detection using similarity profile check	Clone Attack	Suited only for high profile accounts while low profile attacks are difficult to identify
(Wu et al. 2017)[30]	Social Norm Incentives	Sybil attacks in networks	Suitable for small networks but is not suited for complex networks
(Anjos et al. 2014)[2]	Attack detection using face recognition	Photo Attack detection	Used only for photo attack in social media
(Amerini et al. 2011)[1]	Copy move attack	SIFT Based mechanism for attack detection	Can be implied on large image sets but not tested on textual information
(Shi et al. 2017)[24]	Event attack detection	Event detection in social media	Fixed datasets or static datasets uses produce effective results but dynamic datasets still not checked

Table 1: Comparison of attack detection strategies

4. CONCLUSION AND FUTURE SCOPE

From the analysis conducted we conclude that the deception is a common problem in the field of online social media. The steps must be taken in order to prevent the deception. The causes of deception is lack of structure to ensure that only valid users can enter into the system. The proper validation mechanisms are missing since the OSN is typically concerned about the length of the database rather than security of the system. This is a prime factor which is leading to the deception.

In the future some sort of security mechanisms must be enforced to ensure the validity of the user. This can be accomplished by the use of background check mechanisms to prevent clone attacks.

REFERENCES

[1] Amerini, I. et al., 2011. A SIFT-Based Forensic Method for Copy – Move Attack Detection and Transformation Recovery. , 6(3), pp.1099–1110.

[2] Anjos, A., Chakka, M.M. & Marcel, S., 2014. Motion-based counter-measures to photo attacks in face recognition. , (November 2012), pp.147–158.

[3] Aprem, A. & Krishnamurthy, V., 2016. Utility Change Point Detection in Online Social Media : A Revealed Preference Framework. , (c), pp.1–12.

[4] B, J.R. et al., 2016. Detecting Overlapping Community in Social. , pp.99–110.

[5] Barbera, M. V & Mei, A., 2012. Personal Marks and Community Certificates : Detecting Clones in Wireless Mobile Social Networks.

[6] Bhat, S.Y. & Abulaish, M., 2014. Communities A gainst Deception in Online Social Networks 1 The Platform 2 The Mischef. , 2014(2), pp.8–16.

[7] Bu, K. et al., 2015. Deterministic Detection of Cloning Attacks for Anonymous RFID Systems. , 11(6), pp.1255–1266.

[8] Caton, S. et al., 2014. A Social Compute Cloud : Allocating and Sharing Infrastructure Resources via Social Networks. , 1374(c), pp.1–14.

[9] Choi, S., Chung, K. & Yu, H., 2013. Fault tolerance and QoS scheduling using CAN in mobile social cloud computing.

[10] Dave, D., Mishra, N. & Sharma, S., Detection Techniques of Clone Attack on Online Social Networks : Survey and Analysis. Elsevier, pp.179–186.

[11] Devi, J.C. & Poovammal, E., 2016. An Analysis of Overlapping Community Detection Algorithms in Social Networks. Procedia - Procedia Computer Science, 89, pp.349–358. Available at: http://dx.doi.org/10.1016/j.procs.2016.06.082.

[12] Egele, M. et al., 2015. Towards Detecting Compromised Accounts on Social Networks. , 5971(c).

[13] Etter, M. et al., 2017. Measuring Organizational Legitimacy in Social Media : Assessing Citizens ' Judgments With Sentiment Analysis.

[14] Grewal, R. & Scholar, P.G., 2015. A Survey on Proficient Techniques to Mitigate Clone Attack in Wireless Sensor Networks. , pp.1148–1152.

[15] Gupta, S. & Arora, S., A Hybrid Firefly Algorithm and Social Spider Algorithm for Multimodal Function. , pp.17–30.

[16] Hovy, D. & Spruit, S.L., 2001. The Social Impact of Natural Language Processing.

[17] Khabbazian, M., Mercier, H. & Bhargava, V.K., Wormhole Attack in Wireless Ad Hoc Networks : Analysis and Countermeasure.

[18] Kiruthiga, S., 2014. Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques.

[19] Kontaxis, G. et al., Detecting Social Network Profile Cloning.

[20] Maio, C. De et al., 2017. Unfolding social content evolution along time and semantics. Future Generation Computer Systems, 66, pp.146–159. Available at: http://dx.doi.org/10.1016/j.future.2016.05.039.

[21] Mohammed, J. et al., 2014. Internet of Things: Remote Patient Monitoring Using Web Services and Cloud Computing. In 2014 IEEE International Conference on Internet of Things(iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom). IEEE, pp. 256–263. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7059670 [Accessed January 26, 2016].

[22] Ren, Y., Chen, Y. & Chuah, M.C., Social Closeness Based Clone Attack Detection for Mobile Healthcare System.

[23] Rizi, F.S., Khayyambashi, M.R. & Kharaji, M.Y., 2014. A New Approach for Finding Cloned Profiles in Online Social Networks. , 6(April), pp.25–37.

[24] Shi, L. et al., 2017. Event Detection and User Interest Discovering in Social Media Data Streams. , 3536(c).

[25] Solanki, S., 2016. Related Study of Soft Set and Its Application A Review. , 7(4), pp.15–22.

[26] Tekieh, M.H. & Raahemi, B., 2015. Importance of Data Mining in Healthcare. Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 - ASONAM '15, pp.1057–1062.

[27] Tsikerdekis, M. & Ieee, M., 2017. Real-Time Identity Deception Detection Techniques for Social Media : Optimizations and Challenges.

[28] Tsikerdekis, M. & Zeadally, S., 2014. Multiple Account Identity Deception Detection in Social Media Using Nonverbal Behavior. IEEE Transactions on Information Forensics and Security, 9(8), pp.1311–1321. Available at: http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=6843931 [Accessed February 25, 2016].

[29] Wang, W. & Zhang, Y., 2007. On fuzzy cluster validity indices. Fuzzy Sets and Systems, 158(19), pp.2095–2117.

[30] Wu, C., Gerla, M. & Schaar, M. Van Der, 2017. Social Norm Incentives for Network Coding in MANETs. , pp.1–14.

[31] Zhou, H.W.J.L.L., 2011. Lightweight and effective detection scheme for node clone attack in wireless sensor networks. , (December 2010), pp.137–143.