Innovative Science and Technology Publications

International Journal of Future Innovative Science and Technology,
ISSN: 2454- 194X Volume-4, Issue-2, May - 2018

TARGETED REPLICATED ATTACKS IN FORMATION & MANAGEMENT OF SOCIAL NETWORKS USING PERSISTENT THREATS

N.Jayachitra

PG Scholar

Computer Science and Engineering Department,

Vivekanandha College of Engineering for Women,

Namakkal, India.

jaychithu.jc@gmail.com

M.Mailsamy

Assistant Professor

Computer Science and Engineering Department,

Vivekanandha College of Engineering for Women,

Namakkal, India.

mailsamym@gmail.com

May – 2018

www.istpublications.com

Journal of Future Innovative Science and Technology, ISSN: 2454- 194X , Volume-4, Issue-2, MAY - 2018

TARGETED REPLICATED ATTACKS IN FORMATION & MANAGEMENT OF SOCIAL NETWORKS USING PERSISTENT THREATS

N.Jayachitra

PG Scholar
Computer Science and Engineering Department,
Vivekanandha College of Engineering for Women,
Namakkal, India.
jaychithu.jc@gmail.com

M.Mailsamy

Assistant Professor
Computer Science and Engineering Department,
Vivekanandha College of Engineering for Women,
Namakkal, India.
mailsamym@gmail.com

Abstract: Exploration is initial and vital phase of successful advanced persistent threat (APT). In many cases, High level system architecture: main design of Synchronic Trap is clustering analysis. It measures pair wise user behavior similarity and then uses a hierarchical clustering algorithm to group users with similar behavior over a period of time together. The large volume of user action data leads to a low signal-to-noise ratio, making it stiff to achieve immense detection accuracy. The sheer volume of activity data prohibits a practical implementation that can manage with generic actions. To handle massive user activities at Face book-scale online social network (OSN), we apply divide-and-conquer. We slice the computation of user comparison into smaller jobs along the time dimension and use parallelism to scale. The diversity of normal user behavior and the stealthness of malicious activity hinder high accurate detection. In order to achieve high accuracy, we design Synchronic Trap based on our understanding of an attacker's economic constraints. It is challenging to develop a generic solution that can adapt to new applications.

Index Terms: *Wireless networks, SPAM, Social bots Advanced persistent threats (APTs), social Network security.*

1. INTRODUCTION

The traditional security solutions, such as intrusion prevention systems and endpoint protection have failed repeatedly to mitigate such threats. Several recent studies have expressed the need for new tools and methods specifically aimed at detecting APTs. Deploying new and versatile technologies for identifying and investigating suspicious activities is the only way to survive in the cyber arms race. APTs usually follow a methodological multistage process for conducting a cyber attack. The defense approaches focus on later stages of an APT, Synchronic Trap is built on top of Hadoop MapReduce stack at Face book. Clustering module is done on Graph and large graph processing platform based on the Bulk Synchronous Parallel (BSP) model. For a example The graph compares the photo-uploading activities of malicious users to those of normal users at Facebook..

i) Cyber attacks

A cyber attack is any type of offensive manoeuvre employed by nation-states, individuals, groups, or organizations that targets computer information systems, infrastructures, computer networks, and private computer devices by various means of malicious acts usually originating from an anonymous source that either steals, alters, or destroys a specified target by hacking into a susceptible system. These can be labeled as either a cyber campaign, cyber warfare or cyber terrorism in diverse context. Cyber attacks can range from install spyware on a PC to attempts to destroy the infrastructure of whole nations. Cyber attacks have become more and more sophisticated and hazardous as the Stuxnet worm recently demonstrated. User behavior analytics and SIEM are used to end these attacks. Legal experts are seeking to limit use of the term to incidents cause physical damage, unique it from the more schedule data breaches and broader hacking activities.

ii) Syntactic attacks

In detail, there are a numeral techniques to use in cyber-attacks and a variety of ways to manage them with individuals or establishment on a broader scale. Attacks are broken down into two

categories: syntactic attacks and semantic attacks. Syntactic attacks are simple; it is considered malicious software which includes viruses, worms, and Trojan horses.

iii) Semantic attacks

Semantic attack is the alteration and dissemination of correct and incorrect information. Information customized could have been done without the use of computers even though new opportunities can be establish by using them. To set someone into the incorrect direction or to cover your tracks, the dissemination of incorrect information can be utilized.

2. LITERATURE REVIEW

This section gives a detailed review about various deduction schemes. Here we reviewed how the deduction problem is resolved in each scheme. While reviewing a scheme we listed the algorithms and techniques used in that scheme and the merit and demerit of that scheme are also specified. The following papers are survived in this section

In [1] authors Nikos Virvilis, Oscar Serrano Serrano and Bart Vanautgaerden in the paper **"Changing the game: the art of deceiving sophisticated attackers"** has an insider attacks resulted in the ex-filtration of highly confidential information to the public. Traditional security solution has futile repeatedly to moderate such threats. In order to secure against such sophisticated adversaries we need to do over our defences, developing technologies paying attention more on detection than on prevention. The main objective of most of these attacks is the ex-filtration of huge amounts of data. An adversary with sophisticated levels of proficiency and major resources, allowing it through the use of multiple diverse attack vectors (e.g. cyber, physical, and deception) to generate opportunities to achieve its objectives, which are typically to establish and extend its presence within the information technology. The objectives frequently extended over a period of time, by adapting a defenders efforts to resist it, and with determination to uphold the level of interaction needed to execute its objectives. The Research also relies on the principle that anomaly detection is suitable for finding new types of attacks, however it is known that machine learning techniques are finest suited to finding events similar to ones seen earlier. Therefore, these approaches show capable detection possibilities for specific (instruction) data sets, but are subject to serious operational boundaries. Therefore the merits are they focused on techniques that are non-intrusive and that will infrequently result in fake positives. We suggest integrating them into an anomaly-detection system incorporating some additional data sources, such as HR databases (e.g. user data, leave data), access rights matrices.

In [2] authors Ms. Rakhi Radheshyam Patel in the paper "**Review on Detecting APT Malware Infections Based on Traffic Analysis and DNS"** describes the Advanced Persistent Threat attacks are increasing on the internet these days. Unfortunately, they are stiff to detect an APT. It is a nonstop hacking processes and set of stealthy targeting a specific entity with high-value information, such as government, military and the financial industry. The intention of an APT attack is to steal data rather than to cause harm to the organization or network. Once hacking into the network has been achieved, by installing APT malware on the infected machine by attacker. For occurrence, APT malware is, Trojan horse or backdoor, is tailored for firewalls and anti-virus software of the objective network. It is not only used for remotely controlling the compromised machine in the APT attack, but also for stealing sensitive information from infected host over an extended period of time. APT malware is very different from the worms and bots. The primary purpose of APT malware is to distantly control the machines and to take confidential information, rather than to launch denial-of-service attacks, cause damage or send spam emails. DNS is popular for malware to place command and control (C&C) servers. The proposed novel system located at the network point that aims to effectively and detect APT malware infection based on malicious DNS and traffic analysis. To detect suspicious APT malware C&C domains the system uses malicious DNS analysis technique, and then analyzes the traffic of the corresponding suspicious IP using anomaly-based and signature-based detection technology. There are extracted 14 features based on big data to characterize

properties of malware-related DNS. The merit is Malicious DNS analysis is first performed to find out suspicious IP addresses of command and control servers in this approach. IDnS can significantly increase the detecting correctness. The main demerit is the fact IDnS is not high-quality at detecting malware infections that do not rely on domains, such as the Trojan it use the IP address directly to locate the command and control server.

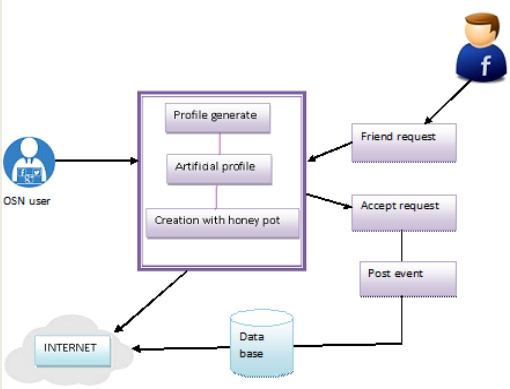


Figure.1 System design and architecture

In [3] authors Twinkle K. Shukla, Dr. D.B.Kshirsagar in the paper "**Machine Learning Approach for Spam Tweets Detection"** propose a method for the massive growth of twitter allows many users to share their information and communicate with each other. Spammers have several goals, which are phishing, advertising, or malware distribution. These goals are comparable to traditional spam in email or blogs, but twitter spam is different. Twitter restrictions the length of each message to less than 140 characters. Because of this limitation, spammers cannot put enough information into each message. To overcome this constraint, spammers usually send a spam containing URLs that are twisted by URL shortening services. When a user clicks the little URLs, he will be redirected to malicious pages. Spam detection built the model which includes the binary classification and these issues is solving by machine learning approach. The machine learning algorithms such as Naïve Bayesian (NB) classifier or Support Vector Machine (SVM) classifier reported the behavior of models. System reported the collision of the data related factors, such as spam to non-spam ratio, timely tweets. The feature of implemented system is easy and time varying spam tweet detection. In this study, System provides a fundamental valuation of ML algorithms on the detection of streaming spam tweets. The merit is system works on offline tweets and real time tweets which are timely updated. System identified that the Feature discretization was an vital pre-process to ML-based spam detection. The System should attempt to bring more discriminative features or better model to further get better spam detection rate.

In [4] authors Luca Maria Aiello, Martina Deplano, Rossano Schifanella, Giancarlo Ruffo in the paper **"People Are Strange When You're a Stranger: Impact and Influence of Bots on Social Networks"** describes a capillary diffusion of online social media and the mounting the number of mobile devices connected to the Internet have determined an unparalleled expansion of the pervasiveness of the Web in real life. Online and offline spheres are today strictly entangled and individual and collective human behavior is increasingly predictable form, or even influenced by, the dynamics of the digital world. Bots are the source of many dangerous attacks or the carrier of unwanted messages, such as spam.Yet, crawlers and software agents are a valuable tool for analysts, and they are continuously executed to gather the data or to examine the distributed applications. However, no one knows which is the actual potential of a bots whose idea is to control a community, to direct consensus, or to influence user behavior. It is commonly assumed that better an agent is to simulates human behavior in a social network, the extra it can be successful to make an impact in that community. We donate to shed light on this concern through an online social experiment aimed to study to what extent a bots with refusal trust, no profile, and no aims to replicate human behavior, can become popular and influential in a social media. We also record that our bots movement unveiled hidden social polarization patterns in the community and triggered an emotional retort of individuals that brings to light subtle privacy hazards perceived by the user base. The merit is we validate the effectiveness of modern link recommendation techniques with an explicit feedback of the user base. Last, we registered profound social dynamics alterations. The demerit is it unveiled Preexistent

hidden social network clusters by triggering the polarization of opinions of community members.

In [5] authors Vinayaka AP, Nalinakshi B G in the paper "**Identifying and Preventing Sybil Attacks in Wireless Networks Using RSSI Value"** is a way to establish communication amongst nodes without having hurdles of wires. Wireless networks may be implemented in a range of manners like mobile ad-hoc networks (MANET), vehicular ad-hoc networks (VANET), wire-less sensor networks etc. mobile ad-hoc network is set of various wireless mobile nodes associated by wireless links. It is an independent system which does not require any pre-existing infrastructure and found for temporary purpose. It is a new technology emerged for fast, in ex-pensive network establishment without having weight of other devices like hub, switches or router. Here, devices are itself capable to behave as node or router to discover a route and forward packets. In MANET, each mobile node has communication range depend upon transmission power, antenna gain and loss along with antenna height. Communication in the networks depends upon the connection among nodes using wireless links. Many defense based on spatial inconsistency of wireless channels survive, but depend either on detailed, multi-tap channel view something not uncovered on commodity 802.11 devices or valid RSSI observations from multiple trusted sources, e.g., corporate access points something not straight accessible in ad hoc and delay-tolerant networks with potentially malicious neighbors. We expand these techniques to be practical for wireless ad hoc networks of commodity 802.11 devices. The merit is to use the inherent difficulty of predicting RSSIs to undo factual and false RSSI observations reported by one-hop neighbors. The demerit is that first it relies on trusted external measurements, e.g., RSSIs from trusted 802.11 access points, which are generally unavailable in open ad hoc networks.

In [6] authors Kyumin Lee, James Caverlee and Steve Webb in the paper "**Uncovering Social Spammers: Social Honeypots Machine Learning"** has a scheme of Web-based social systems permit new community-based opportunities for participants to slot in, share, and interact. This community value and linked services like search and advertising are threatened by spammers, content polluters, and malware disseminators. In an attempt to preserve community value and make sure enduring success, we propose and assess a honey pot-based approach for uncovering social spammers in online social systems. One of the key features of these systems is their reliance on users as primary contributors of content and as annotators and raters of other's content. This reliance on users can lead to numerous positive effects, including large-scale enlargement in the size and content in the community, bottom-up discovery of "citizen-experts", serendipitous discovery of new possessions beyond the degree of the system designers, and new social-based information search and retrieval algorithms. Unfortunately, the relative openness and reliance on users coupled with the prevalent interest and growth of these social systems has also made them key targets of social spammers. Unlike traditional email based spam, social spam frequently contains background information that can enlarge the collision of the spam (e.g., by eliciting a user to click on a phishing link sent from a "friend"). The merit is in general research goal is to explore techniques and enlarge effective tools for automatically detecting and filtering spammers who target social systems with high precisions. The demerit is email spam and phishing approaches relying on data firmness algorithms, machine learning and statistics could inform the further alteration of the proposed approach of low rate of false positives.

In [7] authors S. UmaMaheswari and S.K.Srivatsa in the paper **"Detection of Suspicious URLs Using Real Time System on Social Networks"** has a scheme of Social networking sites are the websites used to communicate and for expressing their interests with others in online. It gives simplicity of contact to new trends /topics and faster communication over longer distances. Some of Internet users use SNS for meeting new friends. Some users use it to locate old friend and relatives. SNS provide users with plenty of benefits like sharing various level of in sequence, media sharing (photo, video, fetch) and many other things. Twitter is one of the eminent social networking and information sharing examines which allows users to fix with worldwide users. When twitter users want to share a URL with

friends via tweets, they usually use URL shortening services to reduce the URL length because tweets can contain only a restricted number of fonts. Malicious users often try to discover a way to attack it. The most common forms of web attacks including spam, scam, phishing, and malware distribution attacks, have also appeared on twitter using URLs. A number of distrustful URL detection schemes have also been introduced. They use static or dynamic crawlers, and they may be executed in virtual machine honey pots such as Capture-HPC and Honey to inspect newly observed URLs. These schemes are unsuccessful beside the feature fabrications or consume much time and resources. We recommend an efficient suspicious URL detection system for twitter. Our system investigates correlations of URL redirect chains extracted from several tweets. Since attackers have partial resources and usually reuse them, their URL redirect chains frequently share the same URLs. The merit is it focuses on the correlations of numerous redirect chains that contribute to the same redirection servers. The demerit is it cannot be deployed as a near real-time system to organize large samples of tweets from the twitter public timeline.

In [8] authors Markus Huber, Martin Mulazzani, Manuel Leithner,Sebastian Schrittwieser and Gilbert Wondracek in the paper "**Social Snapshots: Digital Forensics for Online Social Networks"** proposes that a novel technique for harvesting such statistics from social networking websites. Our approach uses a hybrid system that is based on a routine add-on for social networks in mixture with a web crawling module. The datasets that our tool collects contain profile in order (user data, private messages, photos,etc.) and associated meta-data(internal timestamps and unique identifiers). These social snapshots are significant for security research and in the ground of digital forensics. We implement a prototype for Face book and evaluate our system on a number of human volunteers. We found that these are similar studies, heavily depend on datasets that are collected from the social networking websites themselves, frequently involving data that is harvested from user profiles. In addition, as social networks continue to restore traditional means of digital storage, sharing, and communication, collecting this type of data is also fundamental to the area of digital forensics. In spite of the growing importance of data from OSNs for research, current state of the art methods for data extraction seem to be mainly based on custom web crawlers. The merit is the feasibility and efficiency of our approach and its advantages in contrast to traditional techniques that rely on application specific web crawling and parsing. The demerit is that our techniques cannot be used in cases where no legal assistance with social networking providers exists.

In [9] authors Katharina Krombholz, Heidelinde Hobel, Markus Huber, Edgar Weippl in the paper **"Advanced Social Engineering Attacks"** says about The Internet has develop into the largest communication and information exchange medium. In our daily life, communication has become distributed over a selection of online communication channels. In addition to e-mail and IM communication, Web 2.0 services such as Twitter, Face- book, and other social networking sites have become a part of our daily schedule in private and business communication. This increase in exibility and, conversely, drop in face-to-face communication and shared office space means that increasing amounts of data need to be made available to co-workers during online channels. The development of decentralized data access and cloud services has bring about a paradigm shift in allocation as well as communication, which today is mostly conducted over a third party, be it a social network or any other sort of platform. In global acting companies, teams are no longer geographically co-located, but staffed just-in-time. The decrease in personal interaction combined with a plethora of tackle used for communication (e-mail, IM, Skype, Drop box, LinkedIn, etc.) create new attack vectors for social engineering attacks. Recent attacks on companies such as the New York Times and RSA have uncovered that targeted spear-phishing attacks are an competent, evolutionary stride of social engineering attacks. Combined with zero-day exploits, they become a risky weapon that is often used by advanced persistent threats. The merit of automatic attacks is that the number of probable targets that can be reached within a short period of time is considerably higher than with purely human attacks. The demerit is cannot able to discuss complex advanced attack scenarios.

In [10] authors Tulio C. Alberto, Johannes V. Lochter, Tiago A. Almeida in the paper "**Tube Spam: Comment Spam Filtering on YouTube"** briefly describes about the popularization of broadband about the world has boosted the quantity of Internet users. With faster connections, video host and distribution services became trendy among users. According to a press release of Sandvine, a company focused on standards-compliant network policy control, around 55% of downstream traffic from United States is due to video platforms like Netflix and YouTube. The availability of income through Internet and the broadband connections allowed the appearance of sophisticated new platforms. Recently, YouTube has adopted a monetization system to payment producers, stimulating them to make high quality original content and increasing the amount of visualizations. After the operation of this system, the platform was flooded by undesired content, usually of low quality information known as spam. The profitability promoted by Google in its brand new video distribution platform YouTube has attracted an increasing number of users. However, such success has also attracted malicious users, which aim to self-promote their videos or distribute viruses and malwares. Because YouTube offers inadequate tools for comment moderation, the spam volume is shockingly rising which lead owners of renowned channels to disable the comments section in their videos. Automatic comment spam filtering on YouTube is a challenge even for established classification methods, since the position are very short and often rife with slangs, symbols and abbreviations. The spam found on YouTube is directly related to the attractive profit offered by the monetization system. The merit is Automatic spam filtering on YouTube comments is still an unexplored field, evinced by lack of available tools for comment moderation. The demerit is since they are very short and rife with idioms Regarding the Tube Spam tool, we intend to develop web browser plug-in to filter spam directly from YouTube.

3. PROBLEM STATEMENTS

The honey pots and decoy techniques represent only elementary uses of deception in guard of information systems. This Deception techniques are increases attacker job load and deception allows protector to improved track attack and react before attacker succeed. This technique was tire out attacker's belongings and increases attacker uncertainty. Typically illegitimate will create a honey pot to collect data on individuals and use a variety of Deceptive techniques to steer possible fatalities to the deception.

The deployment of social honey pots for harvesting information of malicious profiles. Analysis of the characteristics of these malicious profiles and those of deployed honey pots for creating classifiers that allows to filter the existing profiles and monitor the new profiles.

To conclude that there is a need to add the artificial profiles to the organization address book and involve HR (and other relevant departments) in order to give the profiles more credibility. In addition, there is a need to increase the employee awareness regarding OSN hazards, and checking profiles of strangers in the organization's address book can be a good solution.

4. PROPOSED STATEMENTS

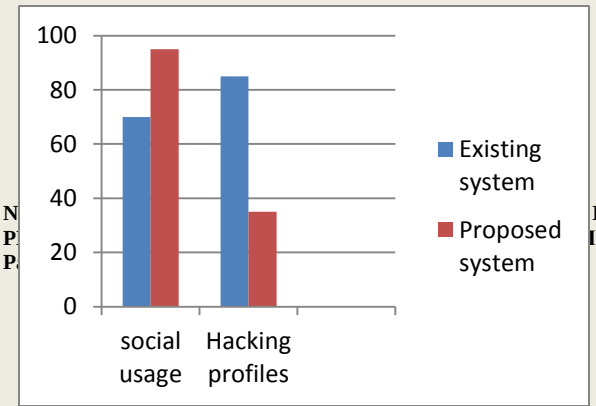
In order to demonstrate the proposed approach and frame- work, we conducted a field trial with the assistance of a large European company (hereafter referred to as Organization A). During eight months, we created, operated, and monitored seven artificial profiles in two OSNs: Xing and LinkedIn. The profiles were generated based on aggregated data provided to us by Organization A, and statistical information we obtained through targeted crawling of the OSNs. It this trial, we limited the number of deployed profiles to seven based on the following reasons: First, each profile was approved by the organization according to a strict selection process that required both time and effort. In addition, incom- ing communication of every profile during the trial had to be inspected by qualified organization employee. Therefore, in order to reduce the inspection effort the company insisted on a "manageable" number of profiles. The main goal of the field trial was to gain insights about the process of creating, integrating, and maintaining social honeypots, a goal that does not necessitate the use of a large number of profiles. Even though,

the goal of this paper was not hunting for attacks, we do believe that even a small number of good and convincing profiles reveal valuable information about the attacks. Based on the knowledge gained in this paper, we plan to conduct a long term, trial with the specific goal of identifying targeted reconnaissance activities.

we used the proposed framework that was described in Section IV, and we used the social network acquisition module to extract statistical information about the organization through targeted crawling of OSNs. Due to the small number of profiles in the case study, the creation of the profile's information was done manually, without the need to use the artificial profile generator module; the profile generation will be described in Section V-C. Following the creation of the profiles, each of the profiles underwent a wiring process using the profile manager module as described in Section IV-C. During the trial, the profiles were monitored by scanning the OSN events and incoming emails using the profile monitor module.

5. PERFORMANCE ANALYSIS

Each honeypot was connected to three types of profiles: Collaborators: Profiles of a few employees within the organization who willingly participated in the trial. Highly connected: Arbitrary popular profiles, which are not affiliated with the organizationand have over 500 connections. Social network profiles that contain the organization name but are not collaborators. We used the profile manager module as presented in Section IV-C to get recommendations about the next friend requests for each artificial profile. An internal investigation was undertaken following a friend request from one of the honeypots to an insider (non-collaborator). The employee searched for the artificial profile and contacted the HR department for information. Due to this investigation, we were requested to stop sending friend requests to insiders after five month since the beginning of the trial.



Aggregated acceptance rate per week for insiders.

The distribution includes the total number of emails, the amount of spam and suspicious emails, emails received from the OSNs, and other messages such as messages from Organization A. We did not obtain USER_6's email data, since Organization A did not forward his/her emails to us. Two of our profiles were exposed to spam. For a period of about a month we registered with several websites (websites offering free services, dating websites, etc.). The purpose of this exposure was to test if it would increase the visibility of the profile and make it more accessible to an attacker, specifically an APT. Exposure to Spam Profiles which were exposed to spam received significantly more spam compared to the other profiles. We noticed that USER_2 received mails from spam websites we did not sign up for. Exposure to spam provides profiles with increased online presence and can improve the legitimacy and attractiveness to the attacker. In addition, USER_2's mail was shared among spam databases.

Challenges:

Challenges During the phase of sending friend requests to insiders, we receivedtwo messages from insiders that replied theycould not find us in the organization's address book. An internal investigation began by the insider. The employee contacted the HR for information. As a result of this investigation, we stopped sending friend requests to insiders. From this event, it is possible to conclude that there is a need to add the artificial profiles to the organization address book and involve HR (and other relevant departments) in order to give the profiles more credibility. In addition, there is a need to increase the employee awareness regarding OSN hazards, and checking profiles of strangers in the organization's address book can be a good solution. The acceptance rate for the profiles of females is higher than that of males, their friend requests are approved at a higher rate. Xing provides information about how active a profile is on the network. The results show that users with higher activity percentages are more

likely to approve incoming friend requests. The acceptance rate of all profiles increased over time, leading to the conclusion that there is a more positive attitude from users toward profiles that have existed for longer periods of time. The results showed that profiles that accepted a friend request from one of our profiles were prone to accepting requests from our other profiles as well. Using this information, it is possible to use an artificial explorer profile with the sole purpose of locating profiles that are likely to approve friend requests, thus increasing our chances for higher acceptance rates.

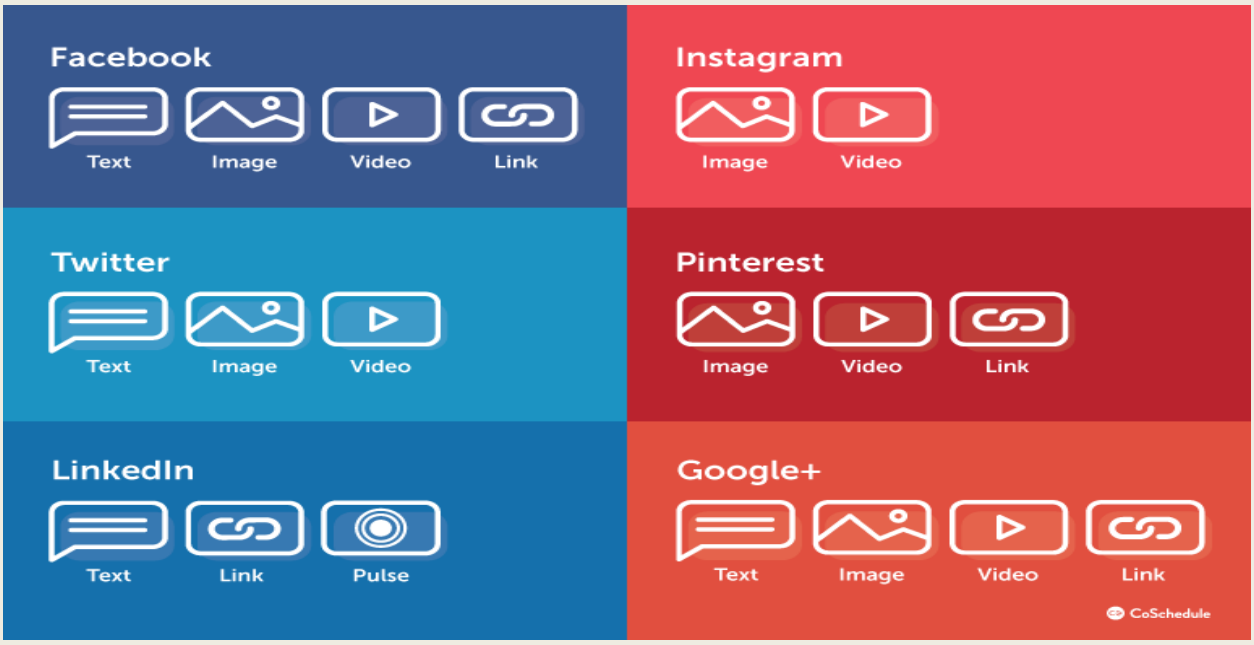
We clustered our conclusions into features inspired in [57]. We defined the following three clusters. Effectiveness: is the measure for deciding whether our framework provides the desired output or not. Being effec- tive means producing the right decision in terms of the emails or friend requests that were identified as suspicious. We used the DCG in order to measure the effectiveness of our framework to detect suspicious activity at an early point in time. Survivability: Reflects how well did the honeypot profiles survive in the social networks and avoided being blocked. Attractiveness: Reflects how genuine and attractive the honeypot profiles were.

6. LEGAL AND ETHICAL CONSIDERATIONS .

Legal Considerations We believe that a minimal violation of terms is the most effective way to reliably estimate the feasibility of an attack and determine how organizations can protect themselves against threats involving social networks. Our framework provides insight into the attacks and greater understanding of the defenses required. Users will benefit from an increase in security by developing a detective and protective mechanism to defend against malicious acts [58]–[60]. The creation of honeypot profiles may lead to the violation of the user terms of a social network, but we believe that our study benefits both users and social network providers. Between 8% and 10% of all social media profiles (approx- imately 150 million profiles) are malicious in nature [10]. This enormous number should emphasize the acute problem that we are facing and demonstrate the need for further study and solutions, particularly since social network providers have repeatedly failed to mitigate such threats. There is a growing need for new tools and methods for detecting such threats, and we believe that it is incumbent on cyber security researchers in academia to address this challenge.

Ethical Considerations Our main goal is to help those targeted by malicious attackers through social networks, while also respecting those individuals. We accomplish this by considering the ethical issues involved with our research and those we are trying to help, and by doing everything in our power to minimize an invasion of their privacy. As part of our effort to respond to ethical considerations, we carried out the following actions (based on the main guidelines in [59]) in order to protect the privacy of the social network profiles, we contacted as follows. 1) The creation and maintenance of the profiles relied upon the cooperation of the organization, and each honeypot profile was created only with the strict approval of the organization. 2) The data used for profile generation were publically available online information. 3) The profiles' privacy settings were defined such that an external entity could not obtain information about the profiles we contacted. 4) We did not access or store any information about the profiles we contacted, and we only recorded the fact of accepting the friend request. 5) The profiles' identifiers were stored securely on a pass- word protected server during the study. 6) Only incoming friend requests and incoming emails were processed and analyzed for the research. 7) The profiles were deleted from the two social networks at the end of the study. 8) We used the OSNs' APIs for the monitoring and crawl- ing process. 9) Identifiers of the profiles that contacted our pro- files or were contacted by our profiles were deleted, and only aggregated statistical data was retained. 10) Data processing and analysis are as follows. a) Email communication with the profiles was inspected and cleared by a security officer of Organization A in order to avoid unintentional leakage of information. b) Suspicious emails that did not contain any per- sonal or confidential information were forwarded to the research team and were inspected by auto- matic tools for determining spam, newsletters, and detecting malicious content/code.

7. SAMPLE OUTPUT



8. CONCLUSION

In this paper, we propose a method that is based on social network honey pots to detect APT attacks at early phases of the APT life cycle. We implemented the method and conducted a field trial to demonstrate the effectiveness of the suggested method with the cooperation of a European organization. The artificial profiles that we created were able to assimilate into OSNs and appeared genuine and attractive to other users. We can conclude that the wiring method proved successful by having more than 70% average acceptance rate when sending friend requests to members of the organization from the artificial profiles. The artificial profiles received suspicious friend requests and emails. We were unable to completely validate the indications of potential forthcoming attacks during our case study. This could be attributed to: the short period of time in which the case study was performed, the small number of profiles we created for the purpose of demonstrating the framework, and/or stopping the wiring process before it finished due to investigation by organization employees. In the future, we plan to continue and scale up the case study for an additional time period in order to further examine the effectiveness and attractiveness of the generated profiles, generating a more diverse pool of profiles, and significantly increasing the number of created profiles. Furthermore, since the number of profiles was small, we created the profile information manually so there was no need to use the complete framework. For the next study, we plan to use and examine the complete framework. We plan to upgrade the automatic generation of the artificial profiles component to support not only the automatic generation of basic profile information, but also the automatic generation of more advanced profile information such as employment and education history.

REFERENCES

[1] N. Virvilis, B. Vanautgaerden, and O. S. Serrano, “Changing the game:The art of deceiving sophisticated attackers,” in Proc. 6th Int. Conf. IEEE Cyber Conflict (CyCon), Jun. 2014, pp. 87–97.

[2] N. Villeneuve and J. Bennett. Detecting Apt Activity With Network Traffic Analysis. Trend Micro Incorporated, accessed on 2012.

[3] Twinkle K. Shukla, Dr. D.B.Kshirsagar, Machine Learning Approach for Spam Tweets Detection, Vol. 5, Issue 5, May 2017

[4] L. M. Aiello, M. Deplano, R. Schifanella, and G. Ruffo, "People are strange when you're a stranger: Impact and influence of bots on social networks," Links, vol. 697, no. 483, pp. 1–566, 2012.

[5] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A nearoptimal social network defense against sybil attacks," in Proc. IEEE Symp. Secur. Privacy, May 2008, pp. 3–17.

[6] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots+ machine learning," in Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2010, pp. 435–442.

[7] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in Proc. 26th Annu. Comput. Secur. Appl. Conf., 2010,pp. 1–9.

[8] A. Paradise, A. Shabtai, and R. Puzis, "Hunting organization-targeted socialbots," in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal.Mining, Aug. 2015, pp. 537–540

[9] Section 9 lab. Automated LinkedIn Social Engineering Attacks, accessed on Sep. 1, 2014. [Online]. Available: https://medium.com/section-9-lab/ automated-linkedin-social-engineering-attacks-1c88573c577e.

[10] D. Wang, D. Irani, and C. Pu, "A social-spam detection framework," in Proc. 8th Annu. Collaboration, Electron. Messaging, Anti-Abuse Spam Conf., 2011, pp. 46-55.