International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 7, pp: 486-490 www.ijaem.net ISSN: 2395-5252

A Survey on Reasoning on Future Cyber-Attacks throughSocio-Technical Hacking Information

Neha Singh nath, Prof. Preeti Ahirwar
M.Tech Scholar (Cyber Security), Assistant Professor (CSE)
VITM, Indore, India.

Date of Submission: 25-09-2020 Date of Acceptance: 08-10-2020

ABSTRACT: The role of computers and the Internet in modern society is well recognized. Recent developments in the fields of networking and cyberspace have greatly benefited mankind, but the rapid growth of cyberspace has also contributed to unethical practices by individuals who are bent on using the technology to exploit others. Such exploitation of cyberspace for the purpose of accessing unauthorized or secure information, spying, disabling of networks and stealing both data and money is termed as cyber attack.With rise in security breaches over the past few years, there has been an increasing need to mine insights from social media platforms to raise alerts of possible attacks in an attempt to defend conflict during competition. The information from the dark web forums is used by leveraging the reply network structure of user interactions with the goal of predicting enterprise cyber attacks. This paper presents a survey on probable cyber security threats through hacking of socio technical data.

Keywords: Cyber Attacks, Cyberspace, Dark Web, Security Breaches, Socio-Technical data.

I. INRRODUCTION

The world is today dominated by technology. Ever since the industrial revolution various new technologies have been developed which have contributed to the improvement of lifestyle. Computers have refined from bulky, complex machines to user friendly and interactive machines which could be used by any person. Coupled with Internet the computers have made communication easier. The role of computers and Internet in modern society is well recognized. The use of Internet has created a virtual area of communication called cyber space where fibre optic cables or wires transmit information to and from the Internet. This space has been increasing steadily in size as more information is fed into it. Cyber space has gradually permeated all aspects of human life such as

- Banking
- Medical
- Education
- Emergency services
- Military etc.

The complexity has also been increasing. Such threats are called cyber attacks. These attacks are used to spread misinformation, cripple. As cyber-crimeabsorbs a huge number of vulnerabilities, patching all softwareflaws is not always an option. As soon as the security teamsmake a progress, there is a whole new batch of vulnerabilitiesthat arise, keeping the defenders continually behind. However, as only a small fraction of vulnerabilities areexploited in real-world attacks mining the interest of malicious hackers can even the odds in this security battle.Thus, organizations are constantly trying to identify whichare the probable IT targets of cybercriminals, although theavailable processes for this task are still not effective.With the widespread of cyber-attack incidents, cyber securityhas become a major concern for organizations.Therefore a comprehensive listingof cyber attacks and classifications of attacks form an important component of cyber security initiatives. The basic categories of attackers is given in figure 1.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 7, pp: 486-490 www.ijaem.net ISSN: 2395-5252

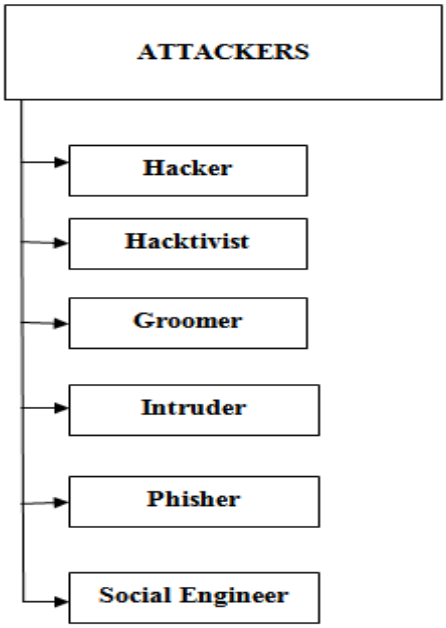

Figure.1 Types of Cyber Attackers and Attacks

II. THE HACKING APPRAOCH BASED ON SOCIO TECHNICAL PLATFORMS

Social hacking describes the act of attempting to manipulate outcomes of social behaviour through orchestrated actions. The general function of social hacking is to gain access to restricted information or to a physical space without proper permission. Most often, social hacking attacks are achieved by impersonating an individual or group who is directly or indirectly known to the victims or by representing an individual or group in a position of authority [1]. This is done through pre-meditated research and planning to gain victims' confidence. Social hackers take great measures to present overtones of familiarity and trustworthiness to elicit confidential or personal information.[2] Social hacking is most commonly associated as a component of "social engineering".Users don't understand or appreciate their vulnerability. For most users, social media offers a means to chat with family and friends, share photos and promote causes.

They underestimate their individual presence and value to outsiders. So, they get careless. After all, if no one out there is really interested, there's no need to worry about passwords, for example among the things they don't understand is the sophistication of the technology. For example, if we download a free game, the game will ask you to connect with friends. To make it easier, the game asks for you to turn over your Facebook friends. Do that and you've established a new network that those friends may, in turn, share with others, and so on.

Preventing hacking is the fundamental principle underlying cyber security initiatives. However,this is essential to consider, especially when devices are connected to the internet. Hackersexploit vulnerabilities,assess weaknesses in Wi-Fi or programs to install a virus, and theninvade the person's privacy through hacking into their internet-connected device. The proactive approach for estimating and mitigating future attacks based on hacker information is given in figure 2.

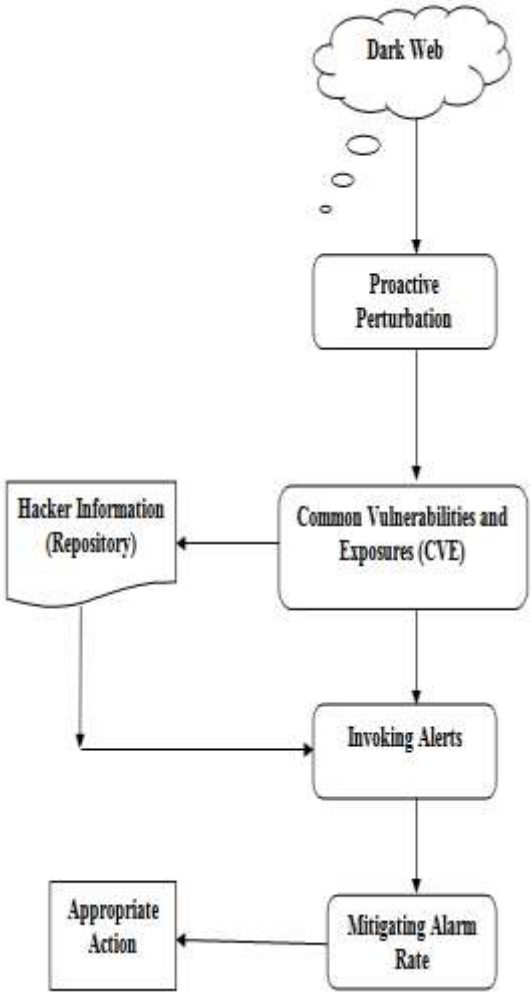

Figure.2 Block Diagram for Mining of Hacker Information to invoke alerts

The above figure depicts the mechanism for proactive reasoning on future cyber-attacks through socio-technical hacker information. In this case, it is assumed that the data on the profiles of hackers on dark web resources can render information about future trends and aspects of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 7, pp: 486-490 www.ijaem.net ISSN: 2395-5252

cyber attacks. The mathematical model of extraction of data from dark web forums is given below:

$$\mathbf{W(v_i, v_j)} = \frac{\mathbf{1}}{\mathbf{M}} \sum_{\mathbf{n}} \forall \mathbf{a, b: V(M, a)} \qquad \mathbf{(1)}$$

$$\mathbf{Or, W(v_i, v_j)} = \mathbf{v_i}^{(\beta\alpha^{(\mathbf{time, M_{k,b}})} - (\mathbf{time, M_{k,a}}))} + \mathbf{V(M_k, b)} \qquad \mathbf{(2)}$$

Here,
F is a dark web forum
W is the correlation between weights
n is the number of threads analysed
v is the number of users posting messages
M is the message number/index
k is the time index
$(M_k, a)\&(M_k, b)$are the messages at time index k for distinct posts a and b in the same thread K.
$\alpha, \beta$are constants with values between 0 and 1.
$v_i, v_j$are distinct messages

Another approach for estimating the similarity co-efficient or the distance among the messages is given mathematically as:
For two lists $\Gamma^1\&\Gamma^2$in the forum 'F', the similarity co-efficient or distance is computed as:

$$\mathbf{D^p(\Gamma^1, \Gamma^2)} = \sum_{\forall\, \mathbf{i,j} \in \mathbf{D(\Gamma^1 \Gamma^2)}} \mathbf{\hat{D}^p_{i,j}\,(\Gamma^1 \Gamma^2)} \qquad \mathbf{(3)}$$

Here,
$\Gamma^1\&\Gamma^2$are two lists
$D^p$is the distance with a penalty p
$\hat{D}^p_{i,j}$takes up fuzzy values for different levels of similarity
(i,j) are the message pair
P is the optimistic penalty parameter

The above distance measure (Kendall's Measure) takes the relative ranking orders of any two elements in the union of two top k lists. Another measure is the absolute distance between the rankings of the same element in the union of two top k lists into consideration called the Spearman's distance measure given mathematically as:

$$F^{k+1}(\mathbf{\Gamma^1, \Gamma^2}) = \sum_{\mathbf{i} \in \mathbf{D_{r1}} \cap \mathbf{D_{r2}}} \left| \mathbf{\Gamma_1'} - \mathbf{\Gamma_2'} \right| \qquad \mathbf{(4)}$$

Here,
F represents the Spearman's distance
$D_{ri}\&D_{r2}$represent the domains of $\mathbf{\Gamma^1}$ and $\mathbf{\Gamma^2}$
$\mathbf{\Gamma_1', \Gamma_2'}$denoted the lists with/without entries in the original lists.
The previous work in the domain is presented subsequently.

III PREVIOUS WORK

This section presents the previous work in the domain with its salient features.

In 2019, Ericsson Marinet al. in [1] proposed Reasoning About Future Cyber-Attacks Through Socio-Technical Hacking Information.With the widespread of cyber-attack incidents, cyber security has become a major concern for organizations. Thewaste of time, money and resources while organizations counterirrelevant cyber threats can turn them into the next victim ofmalicious hackers. In addition, the online hacking community hasgrown rapidly, making the cyber threat landscape hard to keeptrack of. In this work, we describe an AI tool that uses a temporallogical framework to learn rules that correlate malicious hackingactivity with real-world cyber incidents, aiming to leverage these rules for predicting future cyber-attacks. The frameworkconsiders socio-personal and technical indicators of enterpriseattacks, analysing the hackers and their strategies when they are planning cyber offensives online.

In 2018Martin Husák et al. in [2] presented Survey of Attack Projection, Prediction, and Forecasting in Cyber Security .This work provided a survey of prediction, and forecasting methods used in cyber security. Four main tasks are discussed first, attack projection and intention recognition, in which there is a need to predict the next move or the intentions of the attacker, intrusion prediction, in which there is a need to predict upcoming cyber attacks, and network security situation forecasting, in which we project cybersecurity situation in the whole network. Methods and approaches for addressing these tasks often share the theoretical background and are often complementary. In this survey, both methods based on discrete models, such as attack graphs, Bayesian networks, and Markov models, and continuous models, such as time series and grey models, are surveyed, compared, and contrasted.

In 2017Rupinder Paul Khandpur et al. in [3] worked onCrowdsourcing Cyber security: Cyber Attack Detection using Social Media.Social media is often viewed as a sensor into various societal events such as disease outbreaks, protests, and elections. The use of social media is described as a crowdsourced sensor to gain insight into ongoing cyber-attacks. This approach detected a broad range of cyber-attacks (e.g., distributed denial of service (DDoS) attacks, data breaches, and account hijacking) in a weakly supervised manner using just a small set of seed event triggers and requires no training or labeled samples. A new query expansion strategy based on convolution

kernels and dependency parses helps model semantic structure and aids in identifying key event characteristics.

In 2016Aldo Hernández et al. in [4]proposedSecurity attack prediction based on user sentiment analysis of Twitter data. In recent years, security attacks on the web have been perpetrated by hacker activist organizations that aim to destabilize (using different techniques) web services in a specific context for which they are motivated. Predicting these attacks is an important task that helps to consider what actions should be taken if the attack is latent. Although there are applications to detect security threats on the web, currently there is no system that can predict or forecast whether the attacks can reach consummation. This paper presents a sentiment analysis method on Twitter content to predict future attacks on the web. The method is based on the daily collection of tweets from two sets of users; those who use the platform as a means of expression for views on relevant issues, and those who use it to present contents related to security attacks in the web. Daily information is converted into data that can be analysed statistically to predict whether there is a possibility of an attack.

In 2016, Ryan Heartfield et al. in [5] proposed evaluating the reliability of users as human sensors of social media security threats.While the human as a sensor concept has been utilised extensively for the detection of threats to safety and security in physical space, especially in emergency response and crime reporting, the concept is largely unexplored in the area of cyber security. Here, we evaluate the potential of utilising users as human sensors for the detection of cyber threats, specifically on social media. For this, we have conducted an online test and accompanying questionnaire-based survey, which was taken by 4,457 users. The test included eight realistic social media scenarios (four attack and four non-attack) in the form of screenshots, which the participants were asked to categorise as "likely attack" or "likely not attack".

In 2015EktaGandotra et al. in [6] proposed Computational Techniques for Predicting Cyber Threats.With more information becoming widely accessible and new content created every day on today's web, more are turning to harvesting such data and analyzing it to extract insights. But the relevance of such data to see beyond the present is not clear. We present efforts to predict future events based on web intelligence – data harvested from the web – with specific emphasis on social media data and on timed event mentions, thereby quantifying the predictive power of such data. We focus on predicting crowd actions such as large protests and coordinated acts of cyber activism – predicting their occurrence, specific timeframe, and location. Using natural language processing, statements about events are extracted from content collected from hundred of thousands of open content web sources.

In 2015 Nathan Kallus et al. in [7] proposed On the Predictive Power of Web Intelligence and Social Media. With more information becoming widely accessible and new content created every day on today's web, more are turning to harvesting such data and analysing it to extract insights. But the relevance of such data to see beyond the present is not clear. We present efforts to predict future events based on web intelligence – data harvested from the web – with specific emphasis on social media data and on timed event mentions, thereby quantifying the predictive power of such data. We focus on predicting crowd actions such as large protests and coordinated acts of cyber activism – predicting their occurrence, specific timeframe, and location. Using natural language processing, statements about events are extracted from content collected from hundred of thousands of open content web sources.

In 2014William G. Kennedy et al. in [8] proposed Social Computing, Behavioural-Cultural Modelling and Prediction.This research examines the responses of online customers to a publicized information security incident and develops a model of reiterativebehaviours triggered by such a security incident. The model is empirically tested using survey data from 192 users of a recently compromised website. The results of the data analyses suggest that an information security incident can cause a measurable negative impact on customer behaviours, although the impact seems to be largely limited to that particular website. The tested model of reiterativebehaviours indicates that perceived damage and availability of alternative shopping sources can significantly increase reiterativebehaviors of victimized customers, while perceived relative usefulness and ease-of-use of the website show limited effects in reducing such behaviours.

In 2013 Haitao Du et al. in [9] presented Temporal and Spatial Analyses for Large-Scale Cyber Attacks. Prevalent computing devices with networking capabilities have become critical cyber infrastructure for government, industry, academia and every-day life. As their value rises, the motivation driving cyber attacks on this infrastructure has shifted from the pursuit of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 7, pp: 486-490 www.ijaem.net ISSN: 2395-5252

notoriety to the pursuit of profit [1, 2] or political gains, leading to cyber terrorism on various scales.

In 2012 Sun-Hee Lim et al. in [10] presented Prediction model for botnet-based cyber threats.Recent malicious attempts in Cyber-space are intended to emerge cyberwar such as stuxnet as well as to get financial benefits by spam, distributed-of-service(DDoS), identity theft, and phishing through a large pool of comprised hosts, which are called zombies. Botnets are becoming one of the most serious threats to Internet security. We consider that major pre-symptoms of cyber threats are activity and propagation of botnet and propose the prediction model of cyber threats based on botnets.

IV. PERFORMANCE METRICS

The overall performance metrics are mathematically defined as:

Accuracy: It is mathematically defined as:

$$\mathbf{Ac} = \frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}} \quad \mathbf{(5)}$$

Sensitivity: It is mathematically defined as:

$$\mathbf{Se} = \frac{\mathbf{TP}}{\mathbf{TP+FN}} \quad \mathbf{(6)}$$

Recall: It is mathematically defined as:

$$\mathbf{Recall} = \frac{\mathbf{TP}}{\mathbf{TP+FN}} \quad \mathbf{(7)}$$

Precision: It is mathematically defined as:

$$\mathbf{Precisiosn} = \frac{\mathbf{TP}}{\mathbf{TP+FP}} \quad \mathbf{(8)}$$

F-Measure: It is mathematically defined as:

$$\mathbf{F-Measure} = \frac{\mathbf{2.Precision.Recall}}{\mathbf{Precision+Recall}}$$

(9)

Here.

TP represents true positive
TN represents true negative
FP represents false positive
FN represents false negative

V. CONCLUSION:

The cyber security aspect in the current digital space has gained considerable importance. With the advancement in technology, there has also been increase in cyber attacks and security breaches.Using machine learning models to predict security threats has many open research fields including predicting whether vulnerability would be exploited based on Dark Web sources.Hackers can use many sophisticated methods to mine data based on information of users on the social media. Social media analytics is one of the most important forms of data analytics in current scenario of cyber world. This survey presents a comprehensive study on the prediction of cyber attacks based on data from social technical platforms.

REFERENCES

[1]. Ericsson Marin, Mohammed Almukaynizi, Paulo Shakarian, , "Reasoning About Future Cyber-Attacks ThroughSocio-Technical Hacking Information", IEEE 2019

[2]. Martin Husák , Jana Komárková , Elias Bou-Harb , Pavel Čeleda, "Survey of Attack Projection, Prediction, and Forecasting in Cyber Security, San Deago, IEEE 2018

[3]. Rupinder Paul Khandpur, Taoran Ji , Steve Jan , Gang Wang,Chang-Tien Lu, NarenRamakrishnan "Crowdsourcing Cybersecurity: Cyber Attack Detection using Social Media", ACM Digital Library 2017.

[4]. Aldo Hernández , Victor Sanchez , Gabriel Sánchez , Héctor Pérez , Jesús Olivares , Karina Toscano , Mariko Nakano , Victor Martinez, "Security attack prediction based on user sentiment analysis of Twitter data", IEEE 2016

[5]. Ryan Heartfield , George Loukas, "Evaluating the reliability of users as human sensors of social media security threats", IEEE 2016.

[6]. EktaGandotra,Divya Bansal ,Sanjeev Sofat, "Computational Techniques for Predicting Cyber Threats, Springer 2015.

[7]. Nathan Kallus, "On the Predictive Power of Web Intelligence and Social Media," Springer,2015.

[8]. William G. Kennedy, Nitin Agarwal, Shanchieh Jay Yang ,"Social Computing, Behavioral-Cultural Modeling and Prediction", Springer 2014.

[9]. HaitaoDu ,Shanchieh Jay Yang, "Temporal and Spatial Analyses for Large-Scale Cyber Attacks", Springer 2013.

[10]. Sun-HeeLim ,Seunghwan Yun , Jong-Hyun Kim ,Byung-gil Lee, "Prediction model for botnet-based cyber threats", IEEE 2012.