Journal of Engineering Sciences

Vol 11, Issue 7,July/2020
ISSN NO:0377-9254

A Taxonomy of Security Services in Security-attacks through Artificial Intelligence

Chokkam Sravanthi[1], Jetram Lakshmi[2], Golla Venkateswarlu[3]
[1]Assistant Professor, CSE Dept., KMIT, Narayanaguda, Telangana, sravanthi.chokkam7@gmail.com
[2]Assistant Professor, CSE Dept., KMIT, Narayanaguda, Telangana, jetramlaxmi@gmail.com
3Assistant Professor, CSE Dept., KMIT, Narayanaguda, Telangana, gvenkatesh053.kmit@gmail.com

***Abstract*— Security attacks are most common nowadays Security agencies are perennially in a fight against cyber- attacks and are constantly looking for improved technological solutions to help them in their effort. Artificial Intelligence and Machine learning have capabilities to process and analyse huge data for information and also learn from such processing to effectively detect malignant activity and has great scope to support cyber security agencies. Recent trends indicate attack strategies also incorporate AI concepts thus making fighting them more challenging. There is unanimous recognition that a combined effort incorporating benefits of both machine learning and human intelligence will greatly enhance security efforts.**

Keywords— Artificial Intelligence, Machine learning, Cyber Security, Anti-malware

1. INTRODUCTION

Cyber-attacks are on the rise in our daily lives in all spheres of our personal, official and public domain. We live in a hyper-associated world, in which totally everything from banking to amusement to administrative foundation is done on the web. Be it home, organizations, open and private segment associations, government workplaces, aeronautics, transportation, medication, and diversion to give some examples, digital dangers are genuine and harm the wellbeing and security of our lives. In today's world, data protection has become mandatory, for individuals as well as nations. Consequently, the need for cyber security systems and networks have also been increasing thus enhancing their markets.

Hyper-connected workplaces and the growth of cloud and mobile technologies have sparked a chain reaction when it comes to security risks. The vast volume of connected devices feeding into networks provides cyber criminals new and plentiful access points to their targets thus multiplying the threat. This calls for newer technologies to support and enhance the available security control mechanisms [1].

Artificial intelligence (AI) has been an intriguing area of computer science research in recent years. Machine learning (ML) and AI are being applied more comprehensively across businesses and applications than any other time in recent memory as registering power, information assortment and capacity abilities increment AI has the capability to process and analyse the huge data now available and thereby understand new trends and details. This meaningful output can be used by security agencies to effectively combat the threat prevalent in the information network [2].

1.1 TERMINOLOGY AND CONCEPTS

1.1.1 CYBER SECURITY

As a very basic function, the cyber security systems should control viruses and malware in an efficient and effective manner. However, in today's context, the scope of such systems has expanded to a great extent in that they must provide total protection through reliable and efficient systems that warn about imminent danger, provide prevention and protect any kind of important data from malignant attacks [3].

1.1.2 ARTIFICIAL INTELLIGENCE (AI)

Artificial intelligence (AI), in simple terms, is an area of computer science that emphasizes the creation of intelligent machines that work and react like humans. In other words, AI refers to the concepts and development of computing systems which can perform tasks usually associated with human intelligence, like visual insight, speech recognition, decision-making, and translation between languages.

Artificial intelligence is as often as possible related with another term and commonly conversely utilized – Machine Learning (ML). Simulated intelligence is the more extensive idea of machines having the option to complete undertakings such that we would consider

"keen", while ML is based around the possibility that machines are simply offered access to information and let them learn for themselves. The thought is, as opposed to showing the PC severy thing they have to think about the world and how to complete assignments, it may be smarter to instruct them to learn for themselves. ML could be considered as a use of AI [4].

1.1.3 AI IN SECURITY

In modern day technological scenario, computing power, data collection and storage capabilities have increased manifold and widely applied across industries and applications than ever before. AI has capabilities to process and analyse the huge information captured to understand new trends and details. For cyber security, this means new exploits and weaknesses can quickly be identified and analysed to help minimise further attacks. This way, human security personnel can be better utilised. They will definitely be alerted when an action is needed, but they can also spend their time working on more creative, fruitful endeavours [5].

1.2 CURRENT APPLICATIONS OF AI IN SECURITY ATTACKS

Rapid strides have been made in realising the need for a switch from plain rule-based systems to „intelligent" security systems. Consequently, moving away from just being a potential a few years ago, AI based systems have already found its place in several security systems. Some types of such systems are discussed in this section.

1.2.1 DETECTION OF VULNERABILITIES AND SOFTWARE ERRORS

Computer software including those that power our smart devices is subject to coding error, as well as security vulnerabilities that can be exploited by human hackers. Potential consequences can be disastrous, ranging from the safety of an individual to a wider disaster covering an area or a nation. In fact, besides human hackers, experts are now starting to be concerned with attacks by „intelligent" viruses capable of changing behaviour, penetrating targets and modifying drone code. Therefore, systems are being developed and researched that can search out and repair these errors and vulnerabilities, as well as defend against incoming attacks [6]. Globally, several agencies including the military and research universities are funding such efforts

AEG (Automatic Exploit Generation), developed by a start-up "For All Secure" based on research at Carnegie Mellon University in the US, is quoted as the "first end-to- end system for fully automatic exploit generation". AEG works on software deployed on computers and smart devices and can find a bug and determine whether it is exploitable. Once found, it also triggers a defence mechanism which secures vulnerabilities and stops exploitation [7].

Traditional signature-based solutions, such as anti-virus software, are not sufficient in today's threat landscape. Tens of thousands of variants of cyber-attacks are encountered every day, and in such a situation, the time to get protection becomes too long, by which time irreparable damage could have resulted. To effectively shield themselves in real-time, enterprises should use advanced sandboxing security measures, capable of detecting and blocking threats based on dynamic analysis, rather than signatures [8].

1.2.2 SPAM AND PHISHING CONTROL

Before you begin to format your paper, first write and save the content as a separate text file. Complete all substance and authoritative altering before arranging. If you don't mind note segments A-D underneath for more data on editing, spelling and sentence structure.

Machine learning has become a fundamental apparatus in the battles against spam and phishing. Google specialists report that Gmail has utilized ML strategies to channel messages since its dispatch 18 years back. In any case, as attack systems have developed and phishing plans have gotten additionally harming, Gmail and other Google administrations have expected to adjust to programmers who explicitly realize how to defeat them. Regardless of whether attackers are setting up counterfeit Google Docs interfaces or dirtying a spam channel's concept of which messages are malevolent, Google and other enormous specialist organizations have progressively expected to incline toward computerization and ML to keep up [9].

Therefore, Google has discovered applications for ML in a few of its administrations, particularly through a ML strategy known as profound learning, which permits calculations to accomplish progressively autonomous upgrades and self-revamping as they prepare and develop. Earlier, we tended to believe that, if we had more data, we have more problems. Now, with techniques like deep

learning, the more data the better. Google says it is able to prevent violent images, scanning comments, detecting phishing and malware in the Play Store. It is also used to detect fraudulent payments, for protecting the cloud, and detecting compromised computers [10].

1.2.3 POLICING AND CRIME PREVENTION

AI has been in use in Policing for more than 20 years. In early years, software tools were used for predictive policing through a systematic approach that included philosophy and organizational management.

Today, US intelligence agencies use AI along with game theory to predict when terrorists are likely to strike a target.

The US coast guard uses software to generate policy patrol schedules in ports based on passenger load data and traffic changes. Such dynamic schedules make it difficult for a terrorist to predict when there will be increased police presence [11].

1.2.4 PRIVACY PROTECTION

For long, customer data have been used by providers of goods and services to analyse and deliver customized user experience in the product or service. Familiar examples are companies such as Apple, Google and Facebook. However, it is also true that, along with the benefits, comes the security risk for an individual through providing such data. To resolve such increased privacy concerns, the concept of "Differential Privacy" came into being.

Differential privacy offers a method to preserve private data on a network, while providing privacy assurances to the protected subpopulation and using algorithms to investigate the targeted population. Such a solution can be employed when attempting to find patterns or hints of terrorist presence in a civilian population, find infected citizens within a larger healthy population, and other security scenarios. This is however a complex solution since we are dealing with algorithms which are not-yet-discovered. Hence, new routes to violation of privacy may evolve tomorrow and it's much harder to provide, today, technological safeguards against them [12].

2. ML ALGORITHMS AND SECURITY APPLICATIONS

Our principal goal for applying ML to cyber security problems is to find anomalies. More precisely we wish to use it to identify malicious behaviour and segregate this from normal behaviour. However, the biggest challenge is to define what is normal. For example, in between regular internet browsing, occasionally, we download a game or a movie. How does one segregate such a regular activity from a download triggered by some malware? An increase in network traffic might be statistically interesting, but from a security point of view, that does not necessarily represent an attack.

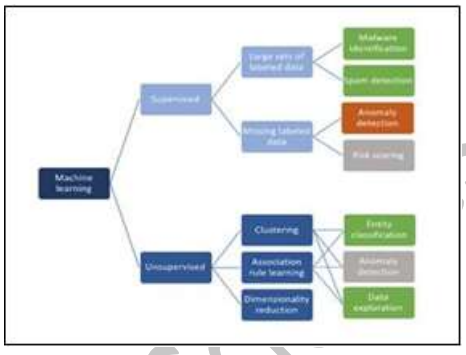

Fig. 2. ML algorithms and security applications- Partial View

Fig: 2 shows a partial view of the ML algorithms and their application for different types of security problems. ML approaches are generally classified into „supervised" and „unsupervised" approaches. These are shown on the left in Fig. 2 and mapped to security problem categories on the extreme right.

The colour of the boxes denoting the problem categories indicate broadly the extent of success achieved. It is seen that, where we have good, labelled data, ML has yielded significant success. These are in cases such as malware identification, spam detection etc. In cases where we don't have good labelled data, though ML has been tried, usable results have not been achieved.

Malware identification is the problem of identifying whether a file is good i.e., whether we can execute it without having to worry about any „side effects"- or whether it is malware that is likely to create a negative impact when we open/run it. Today's approaches in this area have greatly benefited from deep learning where it has helped reduce false positive rates to a manageable level while also simultaneously reducing the false negative rates. Malware identification works so well because of the availability of millions of labelled samples (from both malware and benign applications). These samples enable us to „coach" deep belief networks very well. The problem of spam identification is also very similar to malware identification in the sense that we have a lot of training data to teach our algorithms right from wrong.

But in most other areas, we do not have a lot of data to train our security systems. For example, in the domain of detecting attacks from network traffic. Several attempts have been made over the last two decades to come up with good training data sets for these problems, but we still do not have a suitable one. And without one, training of algorithms is impossible. There are also other problems like the inability to deterministically label data, the challenges associated with cleaning data, or understanding the semantics of a data record.

On the unsupervised side, dimensionality reduction can be taken as an example. Applying it to security data works well, but again, it doesn't really bring us any closer to finding anomalies in our data set. The same is true for association rules. They help us combine similar data records, such as network traffic, but how do we determine inconsistencies with this information? Clustering has the potential to find anomalies – may be to segregate clusters of "normal" and "abnormal" entities, such as users or devices? It was found that the fundamental problems with clustering in security are distance functions and the "explainability" of the clusters.

3. AI IN SECURITY - CHALLENGES

ML tools have shown promising results in providing defence against cyber-attacks already. However, on the flip side, researchers also warn about the ways attackers have begun to adopt ML techniques themselves. Examples already exist. Hacking tools that use machine vision to defeat Captchas have already been reported. And more of these types of attacks are feared to be on the horizon.

Another present threat to ML is data poisoning. If attackers are capable of finding out how an algorithm is set up, or the source of its training data, they can also figure out methods to introduce misleading data that builds a counter- narrative about what content or traffic is legitimate versus malicious. For instance, attackers may run campaigns on thousands of accounts to mark malicious messages or comments as "Not Spam" with the aim of diverting an algorithm's perspective.

Researchers note that this is why it is important that ML systems are set up to encourage human intervention (discussed earlier) so that systems do not work on their own , taking decisions. ML systems should identify deviations from normal behaviour and have the option to refer to a human saying „this is different". In strict terms, there is not much intelligence in here - these are basically, intense processing of data and performing correlations and drawing inferences. It has also been recognised that collaboration among various defenders and with the research community may be necessary to stay ahead of attackers who use ML techniques themselves.

4. CONCLUSION AND FUTURE WORK

For the security industry to get the most out of AI, they need to recognise what machines do best and what people do best. Advances in AI can give new instruments to danger trackers, helping them secure new gadgets and systems even before a danger is grouped by a human analyst.

ML strategies, for example, unsupervised learning and nonstop retraining can keep us in front of the digital hoodlums. In any case, programmers aren't settling for the status quo. We should give our danger analysts the chance to innovatively consider the following attack vector while upgrading their capacities with machines. There is genuine guarantee behind ML in security, in spite of the mind-boggling publicity. The test is holding desires under tight restraints. It merits underlining that AI forecast devices presented with the best possible safeguards and guidelines set up can possibly diminish or evacuate human predisposition instead of validating its belongings.

REFERENCES

[1] Bernard Marr, "What Is The Difference Between Artificial Intelligence And Machine Learning?" Web Article, www.forbes.com, Dec 2016

[2] Hal Lonas, "The role of AI in cyber security" Web Article, www.information-age.com.

[3] Rajat Mohanty, "5 Minute Guide to AI in Cyber Security", Blog www.paladion.net, Nov 2017

[4] Daniel Faggella, "Artificial Intelligence and Security: Current Applications and Tomorrow's Potentials", Web Article, www.techemergence.com, Sep 2017

[5] Lily Hay Newman, "AI can help Cybersecurity – if it can fight through the hype", www.wired.com, Apr 2018

[6] Raffael Marty, "AI and Machine Learning in Cyber Security", www.towardsdatascience.com,

[7] Obanaik, V.: Secure performance enhancing proxy: To ensure end-to- end security and enhance TCP performance over IPv6 wireless networks. Elsevier Computer Networks 50, 2225–2238 (2006)

[8] Bellovin, S.: Probable plaintext cryptanalysis of the IPSecurity

protocols. In: Proceedings of the Symposium on Network and Distributed System Security (February1997) Interworking between IP security and performance

[9] D. Balensonet al., Key management for large dynamic groups: One- way function trees and amortized initialization, IETF Draft, work-in-progress, draft-balenson-groupkeymgmt-oft-00.txt, Feb. 2015.

[10] M. Baugheret al., The group domain of interpretation, IETF Internet Draft, work-in-progress, draft-ietf-msec-gdoi-08.txt, May 2003, expires Nov. 2003.

[11] J. Border et al., Performance enhancing proxies intended to mitigate link-related degradations, IETF, RFC3135, June 2001.

[12] R. Canetti et al., Multicast security: A taxonomy and some efficient constructions, in Proc. IEEE INFOCOM, 1999, pp. 708 – 716.