Countering Cybersecurity Threats with AI

Nemanja Veselinović1,*, Miloš Milašinović1, Miloš Jovanović1, Aca Aleksić2 and Nenad Biga3

1 Belgrade Metropolitan University, Faculty of Information Technology, Tadeuša Kočuška 63, Belgrade, Serbia
2 Academy of Technical and Art Applied Studies, School of Electrical and Computer Engineering, Belgrade, Serbia
3 Graduate School of Business, La Salle University, Philadelphia, United States of America

Abstract
This paper discusses the role of artificial intelligence in cyber attacks. Because of all the specifics that AI technology brings with it, what is characteristic of it is its relationship with cyber risks. AI technology has a twofold effect on cyber security: it can threaten it, and it can also contribute to it. This paper discusses the role of artificial intelligence in cyber attacks. It also talks about machine learning and some of the methods, as well as deep learning methods. Here, various researches are analyzed and the best artificial intelligence methods are presented when it comes to cyber security.

Keywords
AI, Cybersecurity, Machine Learning, SVM, Spam, Threat

1. Introduction
Cyber security is the development of defensive strategies that protect computer assets, networks, data and programs from unauthorized access, alteration or destruction. Due to the great advancement of information and communication technologies, new cyber security threats are emerging and changing rapidly. Cybercriminals are adopting new techniques that make their attacks faster and more extensive.

Thus, there is a demand for more adaptable and compact cyber defense systems that can detect a wide range of threats in real time. In recent years, the adoption of artificial intelligence (AI) techniques has grown and continues to play an essential role in detecting and preventing cyber threats. While the AI agenda was proposed in the 1950s, it has grown rapidly in recent years and now affects all forms of communities and occupations.

This trend also affects the field of cyber security where artificial intelligence is used for both attack and defense in cyberspace. Many fields benefit from artificial intelligence, such as natural language processing, gaming, education, healthcare, manufacturing and more. From an attack perspective, cyber threats can use artificial intelligence to improve the excellence and scale of their attacks.

From a defense point of view, artificial intelligence is used to improve defense strategies, so that defense systems become more flexible and effective, which involves adapting to changes in the environment to reduce the impacts that have occurred.

Recently, researchers have presented several researches in the field of artificial intelligence and cyber security. Some of them only focused on adopting machine learning methods for cyber problems while other research remained focused on deep learning methods. In addition, there is a lack of literature dealing with nefarious uses of AI.

2. The Impact of Artificial Intelligence on Cyber Security
Defining AI can have two approaches. First, it is a science that seeks to develop intelligent machines in which scientists apply information, decision, logic and learning to make machines intelligent and able to think, learn, decide and act while trying to solve a problem, just as it does human reason.

On the other hand, scientists refer to AI as a science that researches and develops methods to solve complex problems that are impossible to solve without adopting intelligence [1].

For example, scientists can build an AI system for real-time analysis and decision-making based on huge amounts of data. In recent years, AI has led to advances in many scientific and technological fields, such as computerized work, natural language processing, expert systems, image recognition, and more.

The rapid development of computer technology and the Internet has a significant impact on people’s daily life and work. However, it has also created many new cybersecurity problems: First, the proliferation of data makes manual analysis impractical [1].

Second, threats are growing at a rapid rate, which also
means that new, short-lived species and highly adaptive threats are becoming quite normal. Third, threats currently threaten various propagation, infection, and evasion techniques; therefore, they are difficult to predict and detect.

It takes a lot of time, money and effort to produce and implement an algorithm. Also, hiring or training people in this field is difficult and expensive. Many deviations and threats occur and continue to spread. So artificial intelligence based methods are expected to keep pace with these cyber security issues [1].

2.1. Positive Use of AI

Based on its large automation and data analysis capabilities, AI can be used to analyze large amounts of data with accuracy, speed and efficiency. An artificial intelligence system can use existing information and recognize threats from the past to identify similar attacks in the future, even if they change. Artificial intelligence has several advantages when it comes to cyber security in the following aspects.

AI can detect new attack changes: Conventional technology mostly relies on known attackers and attacks, while leaving room for blind spots when detecting events in new attacks. The limitations of old defense technology are now being addressed through intelligent technology.

For example, privileged activity on an intranet can be monitored, and any significant mutation in privileged access operations can indicate a potential threat. If the detection is successful, the machine will become more sensitive to detecting similar patterns in the future. With more data, the machine can better learn and adapt to detect faster and more accurate operations [2].

This is very useful as cyber attacks become more sophisticated and hackers develop new and innovative approaches.

AI can handle large volumes of data: AI can improve network security by developing autonomous security systems to detect attacks and respond to breaches. The amount of security alerts that appear every day can be overwhelming for security groups.

Automated threat detection and response has helped reduce the work of network security professionals and can help detect threats more effectively than other methods. When a large amount of data is created and transferred over the network every day, network security professionals will gradually have difficulty tracking and identifying attack factors quickly and reliably [2].

This is where artificial intelligence can help by expanding the monitoring and detection of suspicious activity.

An AI security system can learn to respond better to threats: AI helps detect threats based on application behavior and network-wide activity. An AI security system learns about regular network traffic and behavior and builds a baseline of what is normal. From there, any deviations can be observed to detect attacks.

Artificial intelligence techniques are an emerging area of research that improves security measures for cyberspace.

Many AI methods are used to deal with threats, including intelligent agents, neural networks, computational intelligence, artificial immune systems, data mining, pattern recognition, ML, DL, and others. However, among these techniques, ML and DL have recently attracted much attention and achieved the most achievements in the fight against cyber threats [2].

2.2. Disadvantages and Limitations of Using AI

The advantages highlighted above are only a fraction of how artificial intelligence can help cyber security, but the application of this technology has some limitations.

Datasets: Creating an AI system requires a significant number of input samples, and obtaining and processing the samples can be time-consuming and resource-intensive. Resource Requirements: Building and maintaining the underlying system requires a tremendous amount of resources, including data, memory, and computing power. The qualified resources necessary to implement this technology require significant costs [2].

False alarms: Frequent false alarms are a problem for users because they disrupt business by possibly delaying any necessary response and generally affecting efficiency. The fine-tuning process is a compromise between reducing false alarms and maintaining a level of security.

Attacks on AI-based systems: Attackers can use various attack techniques that target AI systems, such as adversarial inputs, model theft, and data poisoning. One important aspect to consider is the nefarious use of AI. This technology will also be used as a way to improve threats. For example, malicious actors can use ML techniques to generate a variant of malware that is difficult to detect at machine speed. Artificial intelligence could better personalize the phishing scheme and increase the scope of the attack, making the attack more likely to succeed [2].

3. Artificial Intelligence Methodology for Cyber Security

This section provides an overview of learning algorithms, a core concept of AI. In addition, section presents a brief introduction on ML, DL, and bio-inspired computational methods frequently used in the field of cybersecurity.
3.1. Learning Algorithms

AI is a branch of computer science that seeks to produce a new type of intelligent automaton that reacts like human intelligence. In order to achieve this goal, machines must learn. To be more accurate, we need to train the computer using learning algorithms. Learning algorithms help improve task performance through experiential training and learning, so far there are three main types of learning algorithms we use to train machines:

- **Supervised learning**: This type requires a training process with a large dataset that is pre-labeled. These learning algorithms are often used as a classification engine or a regression engine.

- **Unsupervised Learning**: Unlike supervised learning, unsupervised learning algorithms use unlabeled data sets for training. These approaches are often used for data clustering, density estimation, or dimensionality reduction.

- **Reinforcement Learning**: Reinforcement learning is a type of learning algorithm that learns the best actions based on rewards or punishment. Reinforcement learning is useful in situations where data is limited or not provided [3].

3.2. Machine Learning Methods

Machine learning (ML) is a branch of artificial intelligence that aims to strengthen systems by using data to learn and improve without explicit programming. ML has strong ties to mathematical techniques that enable the process of extracting information, drawing conclusions from data, and discovering patterns. There are different types of ML algorithms, but they can be classified into three main categories: unsupervised learning, supervised learning, and assisted learning. In the field of computer security, standard ML algorithms are decision trees (DT), support vector machines (SVM), association rule (AR) algorithms, ensemble learning (EL), k-means clustering and principal component analysis (PCA), Bayesian algorithms, k-nearest neighbor (KNN), random forest (RF) [4].

3.3. Deep Learning Methods

Deep learning (DL) is a subfield of ML and uses data to teach computers how to do things that only humans are capable of. His motivation lies in the working mechanisms of the human brain and neurons for signal processing. The core of deep learning is that if we construct more extensive neural networks and train them with as much data as possible, their performance continues to increase. The most important advantage of DL over conventional ML is its superior performance on large datasets. Similar to ML methods, DL methods also have supervised learning, unsupervised learning, and assisted learning. Typical DL algorithms often used in the cyber security domain are: Feedforward Neural Networks (FNN), Recurrent Neural Networks (RNN), Complex Autoencoders (SAE), Generative Adversarial Networks (GAN), Restricted Boltzmann Machines (RBM), Convolutional Neural networks (CNN), deep belief networks (DBN) and ensemble DL networks (EDLN) [5].

3.4. Bio-inspired Computational Methods

Bio-inspired computing is a branch of AI that has been one of the most studied in recent years. It is a collection of intelligent algorithms and methods that adopt bio-inspired behaviors and features to solve a wide range of complex academic and real-world domain problems.

Among many biologically inspired methods, the following techniques are most commonly used in the cyber security domain: Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Evolutionary Strategies (ES), Genetic Algorithms (GA), and Artificial Immune System (AIS) [6].

4. Artificial Intelligence-Based Approaches to Defense Against Cyberspace Attacks

Recently, scientists have proposed numerous techniques that used AI methods to detect or categorize malware, phishing, network intrusion detection, and spam attacks; counter Advanced Persistent Threat (APT); and identify domains generated by Domain Generation Algorithms (DGA). This part divides the literature into four main groups: malware identification; network intrusion detection; phishing and SPAM identification; and others, which compromises the fight against APT and the identification of DGA.

4.1. Malicious Software Identification

Malware is a general term for many types of malicious software, such as viruses, Trojan horses, exploit worms, retroviruses, botnets, and today, malware is a popular method of cyber-attack. The impact of malware on digital society is very large, so much research has been done on adopting AI techniques to prevent and at least mitigate malware. The latest contributions use intelligence to detect and prevent malware.

The proposed method used logistic regression, support vector machine and random forest classifier and was run on the RIPE benchmark suite for experiments. The authors in [7] reported that the framework has a true positive rate of 99% with a false positive rate of less than 5%.
Meanwhile, scientists have presented a framework for malware classification and detection using data mining and ML classification. In that paper, both signature-based and anomaly-based features for detection were analyzed. Experimental results showed that the proposed method is better than other similar methods.

Another approach used operation codes (OpCode), k-nearest neighbors (KNN), and support vector machine (SVM) as ML classifiers for malware classification. OpCode is represented as a graph and embedded in its own space; then, a single classifier or ensemble of classifiers was used to classify each vector as malicious or benign. The empirical result showed that the proposed model is effective with low false alarm rate and high detection rate [7].

Later, Ie et al. has built a deep learning architecture for intelligent malware detection. In this paper, they used AutoEncoder with Multi-Layer Restricted Boltzmann Machines (RBM) to detect unknown malware [8].

A recent research trend in malware detection is focused on mobile malware in general and Android malware in particular. Machine learning, along with deep learning, has been a significant advance in this field. A deep convolutional neural network (CNN) is adopted for malware identification. The raw sequence of operations from the disassembled program was used to classify the malware. The authors in [9] used a support vector machine (SVM) and the most significant permissions from all permission data to distinguish between benign and malicious applications. The authors presented new ML algorithms, i.e., rotation forest, for malware identity. Artificial Neural Network (ANN) and raw API call sequences are methods for detecting Android malware. A recent study by Wang et al. presented a hybrid model based on deep autoencoder (DAE) and convolutional neural network (CNN) to increase the accuracy and efficiency of large Android malware detection [9].

Another research direction that attracted the attention of scientists was the use of bio-inspired methods for malware classification. These techniques have mainly been used for feature optimization and parameter optimization for classifiers [9].

4.2. Intrusion Detection
An intrusion detection system (IDS) is a system that should protect the system from possible incidents, imminent threats or breaches. Artificial intelligence-based techniques are suitable for IDS development and outperform other techniques due to their flexibility, fast calculations, fast learning and adaptability. Therefore, many researchers have studied intelligent methods to improve the performance of IDS, focused on developing optimized features and improving classifiers to reduce false alarms. Some recent notable studies are listed as follows.

Al-Yaseen et al. [10] combined an extreme learning machine with modified k-means as a model for IDS and a support vector machine (SVM). Using the KDD’99 dataset, their model archived results of up to 95.75% accuracy and 1.87% false alarms. Meanwhile, Kabir et al. introduced a least square support vector machine (LS-SVM) sampling-based method for an intrusion detection system. The proposed methodology was confirmed through the KDD’99 Cup data set and obtained real performance in terms of efficiency [10]. Introduced a fuzzy-based semi-supervised learning approach for IDS. The paper used unlabeled samples with the help of a supervised learning algorithm to improve the performance of the classifier. The algorithm was tested on the KDD’99 Cup dataset and outperformed other benchmark algorithms.

Thing to consider is the use of swarm intelligence (SI) for IDS. Botes et al. [11] presented a new method, namely Ant Miner Classification (ATM), which is a decision tree that uses ACO instead of conventional techniques, such as C4.5 and CART, for intrusion detection. Using the NSL-KDD datasets, their approach achieved an accuracy of 65% and a false alarm rate of 0% [11].

In a recent study by Chen et al. [12] an adaptive coupled multi-level intrusion detection method combining whitelist technology and machine learning is presented. A whitelisted was used to filter communication and a machine learning model was used to identify abnormal communication. In this article, the adaptive PSO algorithm and the artificial fish shoal (AFS) algorithm were used to optimize the parameters for the machine learning model. The method was tested on the KDD’99 Cup, Gas Pipeline and industrial terrain datasets. The empirical result showed that the proposed model is effective in different types of attacks.

A clustering technique based on Fuzzified Cuckoo for anomaly detection is presented. The technique consists of two phases: the detection phase and the training phase. In the detection phase, a fuzzy deterministic approach was used to identify anomalies based on input data and previously calculated distance functions. Experimental results showed that the model was effective with an accuracy rate of 97.77% and a false alarm rate of 1.297%. In the training phase, cuckoo search optimization (CSO), k-means clustering and decision tree criterion (DTC) are combined to estimate the distance functions.

Meanwhile, artificial bee colonies and artificial fish swarm algorithms are included to deal with complex IDS problems. In this paper, a hybrid classification method based on ABC and AFS algorithms is proposed to improve
the accuracy of IDS detection. Datasets NSL-KDD and UNSV-NB15 were used to evaluate the performance of the method [12].

4.3. Phishing and SPAM Detection

A phishing attack is a cyber attack that attempts to steal a user's identity or financial credentials. Today, phishing attacks are one of the most dangerous threats on the Internet. Various new approaches have been used to deal with these problems.

Presented a phishing detection scheme called phishing email detection system (PEDS), which joined neural network development and assisted learning. Their model achieved an accuracy rate of 98.6% and a false positive rate of 1.8%.

Also introduced an anti-phishing method, which used several different ML algorithms and nineteen features to distinguish phishing websites from legitimate ones. The model was found to achieve a 99.39% true positive rate.

Another approach by Feng et al. [13], applied a neural network to identify phishing websites by adopting the Monte Carlo algorithm and the principle of risk minimization. Empirical results showed that their model achieved an accurate detection rate of 97.71% and a false alarm rate of 1.7%.

A recent study he conducted introduced a real-time anti-phishing system that used seven different classification algorithms and features based on natural language processing (NLP). According to the authors, their approach gave a promising result with an accuracy rate of 97.98% [13].

Another study built a stacking model by combining GBDT, KSGBoost, and LightGBM using URL and HTML features to classify phishing websites. The authors reported that their approach achieved an accuracy rate of 98.60% [13].

The term “SPAM” refers to unsolicited e-mail (spam). Spam can lead to inappropriate content and security issues. To overcome the shortcomings of these cyber threats, scientists have recently applied various new, intelligent techniques to build spam filter systems.

A spam categorization technique using modified cuckoo search is designed to improve spam classification. Cuckoo step size search was used for feature extraction and SVM for classification. The proposed approach was tested on two spam datasets—Bare-ling and Lemm-ling—and obtained a competitive result.

Later, the research he conducted proposed a system to filter spam Facebook messages using AI and machine learning based technique. PSO algorithm was adopted for feature selection, and SVM and decision tree for classification.

Recently, Aswani et al. [14] provided a hybrid approach to detect spam profiles on Twitter using social media analytics and bio-inspired computing. Specifically, they used a modified k-means-integrated firefly flight algorithm (LFA) with chaotic maps to identify spammers. A total of 14,235 profiles were used to evaluate the performance of the method. The empirical result showed that the proposed model is effective with an accuracy of 97.98% [14].

A recent study by Faris et al. presented a spam detection and identification system based on genetic algorithm (GA) and random weight network (RVN). According to the experiments, the proposed system achieved outstanding results in terms of accuracy, precision and recall [15].

4.4. Other: Remove APTs and Identify DGA

Some existing works using AI approaches to mitigate other types of cyber threats are presented. More specifically, the methods against APT attacks and DGA are described as follows.

4.4.1. Countering an Advanced, Persistent Threat

An Advanced Persistent Threat (APT) is a sophisticated cyber attack that uses advanced techniques to exploit sensitive data and remain undetected. Attackers often focus on valuable targets, such as agencies of large corporations and government organizations, with the ultimate goal of long-term information theft. To defend against APT attacks, scientists have proposed various artificial intelligence techniques to deal with these cyber threats.

A decision tree was applied to build an IDS to detect APT attacks. It can quickly react to APTs and detect an intrusion early on to minimize damage. Empirical results showed that the proposed system achieved a high APT detection rate.

Meanwhile, Sharma et al. [16] presented a framework architecture for APT detection, which was based on multiple parallel classifiers. According to the authors, the proposed framework achieved high efficiency and accuracy.

Explored how deep neural networks (DNNs), which used raw dynamic analysis features, could be used to attribute APTs to a nation-state. During evaluation with a training set containing 3200 samples, the proposed approach reached an accuracy of 94.6% [16].

Burnap et al. [17] used machine activity metrics and a feature map self-organizing approach to distinguish between legitimate and malicious software. Method has shown promise for APT detection.

Another approach introduced an ML-based approach called MLAPT to identify and predict APTs. According to the authors, their system had the ability to predict APT attacks early. Experiments showed that MLAPT had a
true positive rate and a false positive rate of 81.8% and 4.5% respectively [17].

4.4.2. Identifying Domain Names Generated by DGA

Domain Generation Algorithms (DGA) are algorithms used to create a huge number of pseudo-random domain names to hide the operator’s command and control (C & C) server and avoid detection.

Curtin et al. [18] also applied a similar approach using the generalized likelihood ratio test (GLRT) and obtained promising results.

Yu et al. [19] performed a comparative analysis of architectures based on Convolutional Neural Network (CNN) and Recurrent Neural Networks (RNN), tested using a dataset of one million domain names. The authors reported that all comparative models performed well with high accuracy rates and low false positive rates.

presented a new algorithm based on long-short-term memory (LSTM) network to solve the multi-class imbalance problem in DGA malware detection. Based on the obtained results, the proposed algorithm provided an improvement over the original LSTM.

In a recent study, IF-TF was used for DGA and a machine learning-based hidden channel detection DNS system. the proposed approach achieved an outstanding accuracy of 99.92% [20].

Another approach in proposed a framework for identifying word-based DGAs using word frequency distributions and an ensemble classifier constructed from naive Bayesian, extra-tree, and logistic regression. The authors reported that their method outperformed comparable ones [20].

5. Conclusion

The development of technology greatly facilitates our lives in the future. Every new technology that appears brings with it a huge number of advantages, but unfortunately, the disadvantages are what first catches the eye. People as people, look at everything to abuse. The same situation is with artificial intelligence. The area that gives us so many advantages, is a very big problem for the future because of its shortcomings when it comes to privacy. Fortunately, the power of this technology is so great that it brings with it numerous solutions. The fact is that people are not sufficiently informed about the risk of leaving their personal information on the Internet, regardless of the fact that no one is forcing us to do so. It is very important that the number of people working on data security solutions is higher than those who are trying to abuse it because data protection is the first wall of defense against crime of today.

References

[14] R. Aswani, A. K. Kar, P. Vigneswara Ilavarasan, De-


