FORM 2

THE PATENTS ACT, 1970

[39 of 1970]

&

THE PATENTS RULES, 2003

COMPLETE SPECIFICATION

[*See* section 10 and rule 13]

“A MACHINE LEARNING METHOD TO IDENTIFY AND THWART CYBER-ATTACKS”

Name of the Applicant(s)	Nationality	Address
Panipat Institute of Engineering and Technology, Samalkha	Indian	Principal 70 milestone, GT Road, Samalkha, Panipat, Haryana
Dr. Suresh Chand Gupta	Indian	Professor & Head Computer Science and Engineering Panipat Institute of Engineering and Technology, Samalkha Panipat, Haryana
Dr. Devendra Prasad	Indian	Professor & head Computer Science and Engineering(ET) Panipat Institute of Engineering and Technology Samalkha Panipat
Mr. Puneet Sharma	Indian	Assistant Professor Computer Science and Engineering Panipat Institute of Engineering and Technology Samalkha Panipat
Dr. Manoj Arora	Indian	Professor and Head Electronics and Communication Engineering Panipat Institute of Engineering and Technology Samalkha Panipat
Mr. Amit Kumar Dubey	Indian	Assistant Professor Mechanical Engineering Panipat Institute of Engineering and Technology Samalkha Panipat

PREAMBLE OF THE DESCRIPTION

The following specification particularly describes the invention and the manner in which it is to be performed.

FIELD OF THE INVENTION

The embodiments of the present invention generally relates to the field of cyber-attacks integrated with Artificial Intelligence. More particularly, the present invention relates to a machine learning method to identify and thwart cyber-attacks.

BACKGROUND OF THE INVENTION

The following description of related art is intended to provide background information pertaining to the field of the disclosure. This section may include certain aspects of the art that may be related to various features of the present disclosure. However, it should be appreciated that this section be used only to enhance the understanding of the reader with respect to the present disclosure, and not as admissions of prior art.

While machine learning methods offer promising approaches to identify and thwart cyber-attacks, they are not without their challenges and potential problems. Some of the common issues associated with using machine learning for cyber-attack detection include:

Data Quality and Quantity: Machine learning models heavily rely on high-quality and diverse data for training. Obtaining labeled datasets that encompass a wide range of cyber-attack scenarios can be challenging. Moreover, cyber-attack patterns may change rapidly, and maintaining an up-to-date and representative dataset can be difficult.

Imbalanced Datasets: In cybersecurity, normal network traffic data often significantly outweighs the data related to cyber-attacks. Imbalanced datasets can lead to biased models, where the model might become better at identifying the majority class (normal data) but may struggle with detecting rare cyber-attacks.

Adversarial Attacks: Cybercriminals are well aware of machine learning-based detection systems and may develop adversarial attacks to evade detection. Adversarial attacks involve making small, imperceptible changes to inputs, causing the model to misclassify them.

False Positives and False Negatives: Machine learning models are not perfect and may produce false positives (incorrectly classifying normal behavior as an attack) and false negatives (failing to identify an actual attack). Striking the right balance between detection accuracy and minimizing false alarms is a challenging task.

Model Overfitting and Generalization: Models can overfit to the training data, performing well on known data but failing to generalize to new, unseen data. Ensuring that the model generalizes well to real-world cyber-attacks is crucial.

Interpretability and Explainability: Some machine learning models, such as deep neural networks, can be difficult to interpret and explain. In cybersecurity, explainability is essential to understand why a certain decision was made, especially when the model is used in critical systems.

Runtime Performance: Real-time cyber-attack detection requires models that can process data quickly and efficiently. Complex models with large numbers of features can be computationally expensive, impacting their practicality in real-world applications.

Concept Drift: Cyber-attack patterns may change over time, resulting in concept drift. A model trained on historical data might become less effective in detecting new types of cyber-attacks without continuous updates and retraining.

Data Privacy and Security: Collecting and using data for machine learning may raise privacy and security concerns, especially if the data contains sensitive information about users or organizations.

Dependency on Data Labels: Supervised machine learning methods require labeled data, which can be time-consuming and costly to obtain. Unsupervised or semi-supervised techniques can mitigate this problem to some extent but may not be as effective.

There is therefore a need in the art to provide an efficient machine learning method to identify and thwart cyber-attacks.

OBJECTIVE OF THE INVENTION

Some of the objects of the present disclosure, which at least one embodiment herein satisfies are listed herein below.

The primary objective of the present invention is to provide an efficient machine learning method to identify and thwart cyber-attacks.

SUMMARY OF THE INVENTION

This section is provided to introduce certain objects and aspects of the present disclosure in a simplified form that are further described below in the detailed description. This summary is not intended to identify the key features or the scope of the claimed subject matter.

In an aspect, the present invention generally relates to the field of cyber-attacks integrated with Artificial Intelligence. More particularly, the present invention relates to a machine learning method to identify and thwart cyber-attacks.

BRIEF DESCRIPTION OF DRAWINGS

The accompanying drawings, which are incorporated herein, and constitute a part of this invention, illustrate exemplary embodiments of the disclosed methods and systems in which like reference numerals refer to the same parts throughout the different drawings. Components in the drawings are not necessarily to scale, emphasis instead being placed upon clearly illustrating the principles of the present invention. Some drawings may indicate the components using block diagrams and may not represent the internal circuitry of each component. It will be appreciated by those skilled in the art that invention of such drawings includes the invention of electrical components, electronic components or circuitry commonly used to implement such components.

FIG. 1 illustrates an exemplary architecture in which or with which the present invention identify and thwart cyber-attacks using ML, in accordance with an embodiment of the present disclosure.

DETAIL DESCRIPTION OF THE INVENTION

In the following description, for the purposes of explanation, various specific details are set forth in order to provide a thorough understanding of embodiments of the present disclosure. It will be apparent, however, that embodiments of the present disclosure may be practiced without these specific details. Several features described hereafter can each be used independently of one another or with any combination of other features. An individual feature may not address all of the problems discussed above or might address only some of the problems discussed above. Some of the problems discussed above might not be fully addressed by any of the features described herein.

The ensuing description provides exemplary embodiments only and is not intended to limit the scope, applicability, or configuration of the disclosure. Rather, the ensuing description of the exemplary embodiments will provide those skilled in the art with an enabling description for implementing an exemplary embodiment. It should be understood that various changes may be made in the function and arrangement of elements without departing from the spirit and scope of the disclosure as set forth.

Specific details are given in the following description to provide a thorough understanding of the embodiments. However, it will be understood by one of ordinary skill in the art that the embodiments may be practiced without these specific details. For example, circuits, systems, networks, processes, and other components may be shown as components in block diagram form in order not to obscure the embodiments in unnecessary detail. In other instances, well-known circuits, processes, algorithms, structures, and techniques may be shown without unnecessary detail to avoid obscuring the embodiments.

Also, it is noted that individual embodiments may be described as a process that is depicted as a flowchart, a flow diagram, a data flow diagram, a structure diagram, or a block diagram.

Although a flowchart may describe the operations as a sequential process, many of the operations can be performed in parallel or concurrently. In addition, the order of the operations may be re-arranged. A process is terminated when its operations are completed but could have additional steps not included in a figure. A process may correspond to a method, a function, a procedure, a subroutine, a subprogram, etc. When a process corresponds to a function, its termination can correspond to a return of the function to the calling function or the main function.

The word “exemplary” and/or “demonstrative” is used herein to mean serving as an example, instance, or illustration. For the avoidance of doubt, the subject matter disclosed herein is not limited by such examples. In addition, any aspect or design described herein as “exemplary” and/or “demonstrative” is not necessarily to be construed as preferred or advantageous over other aspects or designs, nor is it meant to preclude equivalent exemplary structures and techniques known to those of ordinary skill in the art. Furthermore, to the extent that the terms “includes,” “has,” “contains,” and other similar words are used in either the detailed description or the claims, such terms are intended to be inclusive in a manner similar to the term “comprising” as an open transition word without precluding any additional or other elements.

Reference throughout this specification to “one embodiment” or “an embodiment” or “an instance” or “one instance” means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present disclosure. Thus, the appearances of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the disclosure. As used herein, the singular forms “a”, “an”, and “the” are intended to include the plural forms as well, unless the context indicates otherwise. It will be further understood that the terms “comprises” and/or “comprising,” when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

The present invention proposes a machine learning method to identify and thwart cyber-attacks. Below are the steps involved:

Data Collection:

The first step in developing a machine learning-based cyber-attack detection system is to collect relevant data. This data could include network traffic logs, system logs, security events, and other relevant information about the network and its users.

Data Pre-processing:

Once the data is collected, it needs to be pre-processed to make it suitable for the machine learning algorithms. This involves data cleaning, normalization, feature extraction, and potentially reducing the dimensionality of the data.

Feature Engineering:

In this step, relevant features are selected or engineered from the preprocessed data. These features serve as input to the machine learning model and are critical for the model's ability to distinguish between normal and malicious behavior.

Training Data Split:

The dataset is split into two parts: a training set and a testing set. The training set is used to train the machine learning model, while the testing set is used to evaluate its performance.

Machine Learning Model Selection:

Various machine learning algorithms can be used for cyber-attack detection, including supervised and unsupervised learning techniques. Commonly used algorithms include decision trees, random forests, support vector machines (SVM), neural networks, and clustering algorithms like k-means.

Model Training:

The selected machine learning model is trained on the training data, where it learns to identify patterns and characteristics associated with normal and malicious behavior.

Model Evaluation:

The trained model is evaluated using the testing set to assess its performance and accuracy in detecting cyber-attacks. The evaluation metrics may include precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

Hyper-parameter Tuning (Optional):

If the selected machine learning model has hyper-parameters, they can be tuned to improve the model's performance further.

Real-time Monitoring:

Once the model is trained and tested, it can be deployed for real-time monitoring of network traffic and systems. It continuously analyzes incoming data and raises alerts when suspicious or malicious activities are detected. Real-time monitoring of network traffic and system activities becomes feasible once the model is fully trained and deployed. By continuously analyzing incoming data, the model efficiently classifies network behaviors as either benign or malicious. In the event of a detected cyber-attack, the model promptly generates alerts, empowering cybersecurity experts to enact prompt mitigation strategies.

Continuous Improvement:

Cyber-attacks and threats evolve over time, so the machine learning model should be continuously monitored and updated to adapt to new attack patterns and techniques. Continuous improvement and adaptability are essential attributes of this machine learning method. As cyber-attacks evolve rapidly, the model undergoes regular updates and retraining, staying abreast of the latest attack patterns and tactics.

In one embodiment, the present invention covers how below methods can be used to identify and thwart cyber-attacks:

Anomaly detection: Anomaly detection is a simple but effective method for identifying cyber-attacks. This method looks for patterns in data that deviate from the norm. For example, if there is a sudden increase in network traffic or a large number of failed login attempts, this could be a sign of a cyber-attack. Anomaly detection can be used to identify a wide variety of cyber-attacks, including DDoS attacks, malware infections, and data breaches.

Machine learning classifiers: Machine learning classifiers can be used to identify malware, phishing emails, and other types of malicious content. These classifiers are trained on a dataset of known malicious and benign samples. When a new sample is presented to the classifier, it will be classified as malicious or benign based on its similarity to the samples in the training dataset. Machine learning classifiers can be very effective at identifying cyber-attacks, but they can also be susceptible to false positives. This means that they may sometimes classify a benign sample as malicious.

Deep learning: Deep learning is a powerful machine learning technique that can be used to identify complex patterns in data. This makes it well-suited for identifying cyber-attacks that would be difficult to identify with traditional machine learning methods. For example, deep learning can be used to identify zero-day attacks, which are attacks that exploit vulnerabilities that are not yet known to the public. Deep learning is still a relatively new technology, but it is rapidly becoming a valuable tool for cybersecurity.

In addition to these methods, machine learning is also being used to develop new security tools, such as intrusion detection systems (IDS) and firewalls. These tools use machine learning to analyze network traffic and identify malicious activity. Machine learning is also being used to develop new security training programs for employees. These programs use machine learning to identify the most common cyber-attack techniques and train employees on how to defend against them.

In yet another embodiment, the future of machine learning methods to identify and thwart cyber-attacks is promising and will likely be influenced by several key trends and developments. As technology advances and cyber threats continue to evolve, machine learning will play an increasingly crucial role in enhancing cybersecurity measures. Here are some potential future directions for this field:

Improved Detection Accuracy: Ongoing research and advancements in machine learning algorithms will lead to improved detection accuracy. Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), will be further refined to better identify complex and subtle cyber-attack patterns.

Real-time Threat Analysis: Future machine learning models will be designed to analyze and respond to cyber threats in real-time, minimizing response times and reducing the potential damage caused by attacks.

Explainable AI in Cybersecurity: As the deployment of machine learning models in critical systems becomes more prevalent, there will be a growing demand for explainable AI. Efforts will be made to enhance the interpretability of machine learning models in cybersecurity to understand how decisions are made and improve trust in the systems.

Integration with Threat Intelligence: Machine learning methods will be integrated with threat intelligence feeds and databases to stay up-to-date with the latest cyber-attack techniques and patterns. This integration will help improve the detection of novel and zero-day attacks.

Adversarial Defense Strategies: The development of robust machine learning models that can withstand adversarial attacks will be a priority. Adversarial defense techniques, such as adversarial training and robust optimization, will be integrated to make models more resilient against evasion attempts.

Federated Learning for Privacy: To address privacy concerns, federated learning approaches will be explored, allowing models to be trained across multiple devices or organizations without sharing sensitive data directly.

Context-aware Cybersecurity: Machine learning models will be enhanced to consider the context in which network traffic and behaviors occur. This context-aware approach will improve the accuracy of attack detection and reduce false positives.

AutoML for Cybersecurity: Automated Machine Learning (AutoML) will gain traction in cybersecurity, enabling cybersecurity professionals with limited machine learning expertise to develop effective detection models.

Cyber Threat Hunting with ML: Machine learning will be used to aid cyber threat hunting, allowing security analysts to proactively search for potential threats within their networks.

Collaborative Defense Systems: Collaborative defense systems, where multiple organizations or entities share threat intelligence and coordinate responses, will leverage machine learning to build more effective collective defenses against cyber threats.

Quantum Computing Impact: The development of quantum-safe machine learning algorithms will become essential in the face of potential threats posed by quantum computing to current cryptographic methods.

In yet another embodiment, the presented machine learning method serves as a powerful ally in the pursuit of securing digital infrastructures from relentless cyber threats. With its ability to accurately identify and thwart cyber-attacks, it offers an indispensable resource to safeguard networks and systems from malicious intrusions, fostering a safer and more secure cyberspace for individuals, businesses, and societies at large.

While considerable emphasis has been placed herein on the preferred embodiments, it will be appreciated that many embodiments can be made and that many changes can be made in the preferred embodiments without departing from the principles of the invention. These and other changes in the preferred embodiments of the invention will be apparent to those skilled in the art from the disclosure herein, whereby it is to be distinctly understood that the foregoing descriptive matter to be implemented merely as illustrative of the invention and not as limitation.

We claim(s)

1. A machine learning method for identifying cyber-attacks, comprising:

   (a) Receiving network traffic data from a network;

   (b) Preprocessing the network traffic data to extract relevant features;

   (c) Training a machine learning model using the preprocessed network traffic data to classify normal and malicious network behavior; and

   (d) Generating an alert when the machine learning model detects a cyber-attack based on the received network traffic data.

2. The method of claim 1, wherein the machine learning model comprises a convolutional neural network (CNN) or a recurrent neural network (RNN).

3. The method of claim 1, wherein the preprocessing step includes filtering and normalizing the network traffic data.

4. The method of claim 1, wherein the generated alert includes information about the type and severity of the detected cyber-attack.

5. A computer-readable medium storing instructions that, when executed by a processor, cause the processor to perform the method of claim 1.

ABSTRACT

A MACHINE LEARNING METHOD TO IDENTIFY AND THWART CYBER-ATTACKS

The present invention summarizes a novel machine learning method designed to effectively identify and prevent cyber-attacks. Leveraging supervised learning with deep neural networks, the method extracts essential features from network data to distinguish normal behavior from malicious activities. Real-time monitoring enables timely detection and immediate response to cyber threats, while continuous model updates ensure adaptability to evolving attack patterns. The method represents a promising advancement in cybersecurity, offering enhanced protection against the ever-growing menace of cyber-attacks, safeguarding digital assets and ensuring a secure digital landscape.