**A hybrid machine learning approach for malicious behaviour detection and recognition in cloud computing**

BY*- Mahdi Rabbani ,Yong Li Wang , Reza Khoshkangini , Hamed Jelodar , Ruxin Zhao, Peng Hu*

**SUMMARY OF RESEARCH PAPER**

BY UTSAV AGRAWAL

**1.INTRODUCTION**

-Cyberattacks and emerging security threats are increasing in prominence daily, penetrating networks, influencing major elections, affecting servers and crippling businesses and financial accounts overnight

-Identifying malicious behaviour at level other than log in level.

-Normal and malicious behaviour are by accumulated packets through network. Using two modules’ *data pre-processing* and *recognition.*

-Proposed approach evaluated on UNSW-NB15 dataset, which was captured from thousands of users in real-time network traffic.

**2.RELATED WORKS**

There is a challenge of making a scalable, adaptable, and lightweight abnormal behaviour detection scheme. A typical malicious behaviour recognition system consists of four modules: *a data source, data pre-processing, a recognition engine,* and *a defence response*.

-A research used Geometric Area Analysis method for detecting anomaly patterns by using Trapezoidal Area Estimation. It worked by calculating the distance between observations.

-In another research, the researchers were able to decrease storage costs by approximately 65% , with a secure encrypted Electronic Medical Records (EMRs) deduplication scheme for Cloud assisted eHealth systems (HealthDep). Another research conducted, used blockchain technology to check weather auditors perform the verification at the prescribed time.

-A research in 2016 proposed Least Squares Support Vector Machine-based IDS (LSSVM) for intrusion detection. In a different research, Multivariate correlation analysis (MCA) was used to analyse network traffic and detect potential DoS attacks. They used triangle-area-based method and achieved high-performance results for DoS detection.

-Other methods proposed are: Artificial immune system (AIS) for anomaly detection, Computer vision techniques (CVT) for DoS detection in which network traffic was treated as an image, triangle area nearest neighbours (TANN) for attack detection using k-means clustering and k-NN classifiers.

-Excellent datasets can be collected by feature selecting behavioural data transformed into digital values using techniques like N-grams or BRO-IDS. Using statistical techniques, a probabilistic discriminative model can be developed for android malware detection.

-Dependency graphs, called Droid SIFT were used to tackle zero-day android malware by proposing an android malware classification system.

-A signature-based intrusion detection system alone is insufficient to deal with modern malware. An intelligent machine learning system is required for modern malware detection. The authors listed the main gaps and limitations identified in related works:

* Lack of self-optimization in learning, self-configuration and preventive capabilities
* Lack of collaboration between detectors
* Good approaches only applicable to specific attacks
* Recognition of zero-day attacks

**3.BACKGROUND**

*3.1 Evidence and observations in cloud-based environments*

- Representing the behaviour in a standard format that is understandable to a machine is a challenging task. EVIDENCE < SUB- BEHAVIOR < BEHAVIOR.

*3.2 Intrusion detection system (IDS)*

-Two types: Misuse direction using a dataset and Anomaly detection by creating a normal profile.

- Famous techniques: Support vector machine (SVM) and artificial neural network (ANN).

- are used to monitor host and network-based environments.

- are deployed and installed to various systems and applications such as, data centres, access points,

Cloud and Internet of Things (IoT).

-Network based IDS: these types of IDS are strategically positioned in a network to detect any attack

on the hosts of that network.

- Host based IDS: they are installed in a host and they can monitor traffics that are originating and coming to that particular hosts only.

*3.3 Normal behaviours*

- Normal behaviour is assigned to users who have legitimate access to a network and engage in normal activities during utilization of the network.

-Two vectors to distinguish normal behaviour from malicious.

*3.4 Contemporary network abnormalities*

- Malicious behaviour categorised by different types of attacks such as Analysis, Backdoors, DOS, Exploits, Fuzzers, Generic, Reconnaissance, Shellcode and Worms.

**4.FRAMEWORK ARCHITECTURE**

It consists of two main modules, *Data pre-processing* and *Recognition***.**

*4.1 Data pre-processing module*

Pre-processing filters network data by removing irrelevant or noisy information that affects the performance of the decision engine for recognizing malicious behaviours. Thereafter, the extracted features are compatible inputs for the recognition system.

-The pre-processing involves four steps: feature creation, reduction, conversion and feature normalization.

4.1.1 Feature Creation

-Network traffic features are captured from raw network packets using tools such as Argus, BRO-IDS and Netmate.

-UNSW-NB15 datasets have been categorized into five different groups to determine the potential characteristics of the user behaviours, these groups include; flow features, basic features, content features, time features, generated features and additional features. These features are established using both transactional flow identifiers (i.e., source and destination IP addresses) and transactional connection times (e.g., 10 or 100 connections per second).

4.1.2 Feature Reduction

-In this step, unimportant and noisy features should be removed.

-Network packets contain important information that might be used to identify malicious behaviours, they should be carefully analysed to separate only the relevant information that helps the recognition phase correctly detect malicious instances.

-We use PCA to arrange a set of variables (features) based on the highest variance for each feature and create a new low-dimensional space of uncorrelated features by removing low-variance variables.

4.1.3 Feature Conversion

Since recognition module can process only quantitative data, a unified format is used to convert all non-quantitative features into numeric ones.

4.1.4 Feature Normalisation

This step is used to scale the feature values into a specific confidence interval. The main benefit of this step is that it removes the bias from the raw data without amending the statistical characteristics of the features.

*4.2 Recognition module*

- It models user’s activities and uses obtained information to predict whether user is normal or malicious.

4.2.1 Training phase

-Only a Single activated neuron will pass 1 whereas other Deactivated neurons will pass 0.

-Using PSO to render fast PNN’s as compared to feed-forward neural networks.

-The network trains using patterns with pre-defined labelled classes in dataset. These help system learn in training phase while improving classification capabilities in testing phase.

-Main layers of PNN-

* Input layer: The first layer of PNN structure is identical to the number of features in all classes that are used to model the user observations. It is necessary to feed all the nodes in the input layer with the equal numerical values from the feature vectors which nourishes it and passes to second layer
* Pattern layer: All neurons in the pattern layer are distributed into k groups (one for each class).
* Summation layer: This layer calculates the conditional functions for class probability approximation using a combination of the previously obtained densities; t
* Output layer: The last stage of the training process is the output layer; in this layer, all neurons have been processed to predict the result.

4.2.2 Particle swarm optimization phase

-Particle swarm optimization phase It is an algorithm which helps in the determination of sigma in PNN systems. It is initialized with a swarm value of 𝜎, and the optimization value is computed for a specific pattern. This makes the structure of PNN a self-adaptive network. Each particle changes and updates its position for every movement until it finds the best position.

4.2.3 Stability of constriction factor

-An initial value close to 1 that gradually decreases to zero is considered a good configuration

-The key point for optimizing the PNN is to find the best spread parameters as the weight of hidden layer. Each optimized PNN classifier will output one probabilistic value. Hence, the average value is obtained by k probabilistic values (P1, P2,…, Pk). Finally, the average value determines the class of every input pattern.

**5.EXPERIMENTAL EVALUATION**

*5.1 Study setup and dataset*

-UNSW-NB15 dataset includes normal observations and 9 families of attack.

-Categorised by 47 features which are divided into five categories: flow, basic, content, time and generated features.

- The IXIA traffic generator is connected to 3 different servers; servers 1 and 3 are used to generate normal behaviours, and server 2 is used to generate malicious behaviours. All servers are connected to two routers; pcap files are collected using router 1 directly and router 2 via firewall.

*5.2 Experimental Results*

5.2.1 Evaluation Objectives

• RQ1: **detection** of normal and malicious behaviours. At this level administrator can take action against those who generate malicious behaviour.

• RQ2: **recognition** of different types of contemporary network attacks (malicious behaviours) as a multi-class classification problem. Classify based on attacks types.

5.2.2 Evaluation metrics

* True positive (TP) is the number of actual malicious evidences classified as types of attacks.
* True negative (TN) is the number of actual malicious evidences classified as normal.
* False negative (FN) is the number of actual normal user classified as normal.
* False positive (FP) is the number of actual normal user classified as types of attack.
* True Positive Rate (TPR), or Recall, is the fraction of malicious evidences correctly classified as attacks; it is computed using the following formula: TPR = Recall = (TP)∕((TP)+(FN))
* False Positive Rate (FPR), is the fraction of normal behaviour patterns that are incorrectly classified as a type of attacks. FPR = (FP)∕((TN)+(FP))
* Precision, is the fraction of patterns that are recognized correctly = (TP)∕((TP)+(FP))
* F-measure, the harmonious mean of precision and recall.

Fmeasure = (2 ∗ Precision ∗ Recall) ∕ (Precision+ Recall)

5.2.3 Evaluation results

- For RQ1 we took testing dataset containing 1000 normal and malicious patterns each.

- TPR 0.98 and FPR 0.02

- For RQ2 900 malicious behaviours used as test set.

-TPR 0.96 and FPR 0.04

*5.3 Comparative study*

- The FSVM method showed higher detection rates for DoS and Probe attacks.

- MCA, EDM, CVT and TANN techniques also performed well only for DoS attacks detection.

- Ultimately PSO-PNN can achieve better performance than other techniques for malicious

behaviours in real-time network traffic and can efficiently discriminate between normal and

malicious behaviours.

**6.DISCUSSION AND CONCLUSION**

**-**Monitor and assess user activities within system and represents user behaviour into understandable format for a learning machine.

-PNN used to train the network and to allow it to learn from patterns of activity associated with different types of attacks and PSO-PNN to optimize the performance of the system and reduce the classification errors.

KEY ASPECTS-

-RQ1 shows 4% error where malicious behaviours classify as normal.

- RQ2 shows proposed system is not only able to detect normal and abnormal behaviour but that it is even capable of classifying different types of malicious behaviours with high performance.