Intrusion detection AI

Anomaly detection systems analyse divergence from normal pattern and watchful to unknown behaviour or novel attacks without any knowledge about them. It causes a higher percentage of incorrect alarms, and has the ability to detect unrecognized attacks.

Misuse detection is based on information from past intrusion pattern. This method is very accurate and efficient to detect an intrusion. However, the major weakness is its inability to detect novel attacks.

NS2 – network simulator

These datasets have been employed by many researchers in this area and were provided by MIT Lincoln Labs.

The main dataset two-hundred thousand data connection.  Divided into two groups for training and testing. Each recorded connection has forty-one features including a class label indicating whether the connection is an attack or normal. Only twenty-four distinctive features from the total forty-one were investigated in this study. This is because some of the features were indifferent for both normal and attack connections. Therefore, these irrelevant features were omitted from the experiments.

this study focuses on eight distinctive features to classify the intrusion, which represent duration, type of protocol, type of service, flag, source bytes, number of access files, number of outbound cmds and service difference host rate.

In AIS the generation of the initial population is very crucial as it will determine the quality of the final results.

6 different models for the 6 different attacks. Measure their effectiveness with A mathematical equation with black letters

Description automatically generated

Where A and B is number of attacks in dataset and number of normal connections and a and b is number of attacks detected and normal connections identifies. Closer to 1 value the better. A diagram of a medical procedure

Description automatically generatedA text on a page

Description automatically generated

284948 connection data,  28494 for testing, 256454 for training. 96% accuracy. Probability of dataset choice for testing effected results, quality of samples and similarity to training material.

<https://ieeexplore.ieee.org/document/8463274>

Initially, primary focus on applying machine learning models, such as decision trees (DTs) and support vector machines (SVMs) to existing intrusion detection systems, and it has now been extended to deep learning approaches, such as convolutional neural networks (CNNs), long short-term memory (LSTM), and autoencoders.

most network flow data is normal traffic, malicious behaviour occurs rarely. Moreover, within malicious behaviour most data are well-known attacks, specific attacks extremely rare. Due to this data imbalance problem, AI models deployed cannot sufficiently learn the characteristics of specific network threats, this may leave the network vulnerable to attacks owing to poor detection performance.

GAN

4 stages

* 1) preprocessing – raw dataset changed to format that deep learning can learn
* 2) generative model training – trained to gen scares data
* 3) autoencoder training – trained through generative model outputs as a feature extraction
* 4) predictive model training. – trained through outputs of the gen model and autoencoder

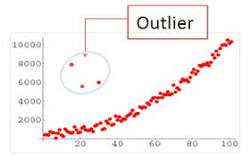
consider three deep learning models that have been widely utilized in AI-based NIDS: 1) deep neural networks (DNNs); 2) CNNs; and 3) LSTM model.

A diagram of a data processing process

Description automatically generated

A white background with black text

Description automatically generated

Outliers are deviations in the dataset, things that stick out.https://www.smarten.com/blog/outlier-analysis-improve-analysis/

Can effect training so they are eliminated. Typically detected by quantifying statistical distribution of the data sets via robust measures of scale. Several standard robust measures of scale for detecting outliers, such as interquartile range (IQR) and median absolute deviation (MAD). Among these measures, we leveraged the MAD.

A math equations and formulas

Description automatically generated with medium confidence

Note that outlier removal should be performed before scaling features, as it can potentially obscure information about outliers.

After outliers, the nominal attributes are turned into one-hot vectors. Each nominal (categorical) attribute is represented as binary vector. For example, “protocol” attribute (commonly in traffic data) with the values tcp, udp, and icmp, attribute is transformed into a binary vector of length 3, and converted into [1, 0, 0], [0, 1, 0], and [0, 0, 1]. With the one-hot encoding process, the numeric attributes are scaled. In general, normalization and standardization are scaling for numeric features. We adopted the min-max normalization method.

the model with a feature extraction did not show significant improvement compared with the model without feature extraction.

splits the given data set according to the classes and then builds generative models for each split subdata set. There are as many gen models as there are classes.

 Note that the trained encoder is placed at the forefront (input layer) of the detection models as a feature extractor and is set not to learn any more when training detection models

A text on a page

Description automatically generated

A diagram of a diagram of a television channel

Description automatically generated with medium confidence

A diagram of a computer component

Description automatically generated

Three main datasets widely used as bench marks and other data from enterprise networks.

NSL-KDD data set

* KDDTrain and KDDTest, with 125973 and 22544 rows
* 41 attributes (3 nominal, 6 binary, and 32 numeric attributes) for different features of the traffic and a label indicating an attack type or normal behavior.
  + 1) Denial of Service (DoS)
  + 2) Probing
  + 3) Remote to Local (R2L)
  + 4) User to Root (U2R).

A table with numbers and a number of objects

Description automatically generated with medium confidence

**UNSW-NB15 Data Set**

* created by the IXIA PerfectStorm tool
* UNSW-NB15\_training and UNSW-NB15\_testing, with 175341 and 82332 records
* 43 attributes 4 nominal, 2 binary, and 36 numeric attributes), whether or not the record is normal traffic (binary-valued attribute) type of attack (when the record is abnormal)
* nine distinct attacks
* Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms.

A table of data with numbers and a number of text

Description automatically generated with medium confidence

**IoT Data Set**

* Mirai botnet scenario
* 23145 IoT traffic
  + 1) Benign – normal traffic
  + 2) C&C
  + 3) DDos
  + 4) PortScan
* 21 attributes (11 nominal, 2 binary, and 8 numeric attributes) 4 removed (ID and IP for eg)
* 100000 Benign data, 6706 C&C data, 14394 DDos data, and 122 PortScan data.

**Real Data Set**

* 4552316 data were labeled as “Normal,” and 230026 data were labeled as “Threat”
* 16 basic features

**A screenshot of a computer

Description automatically generated**

**A text on a page

Description automatically generated**

**DNN, CNN and LSTM of 2 layer**

DNN, first hidden layer 32 neurons and second 16 neurons.

CNN, 1-D-CNN with two convolutional layers. 32 filters, window size of 5. Fully connected layer of 16 neurons follows. Max pooling layer with window size 3 for first layer. Batch normalization layer after each convolutional layer. ReLu in activation function.

LSTM, 64 LSTM cells in each layer and concatenated a fully connected layer with 32 neurons.

default number epoch to 300 and applied the early stop technique (stop learning when relative differences of loss are less than 10-6 consecutively for 35 epochs

four metrics to evaluate the performance of AI models: Accuracy, Precision, Recall, and F1 -score.

A table of numbers and a few percentages

Description automatically generated with medium confidence

A graph of different colored lines

Description automatically generated

For threat type classification included.

A table of data with numbers and symbols

Description automatically generated with medium confidence

<https://ieeexplore.ieee.org/document/9908159>

<https://www.kaggle.com/models?query=biggan&tfhub-redirect=true>

<https://github.com/artcg/BEGAN>

A diagram of a computer network

Description automatically generated

<https://www.youtube.com/watch?v=_gHMkEKGwBM>

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

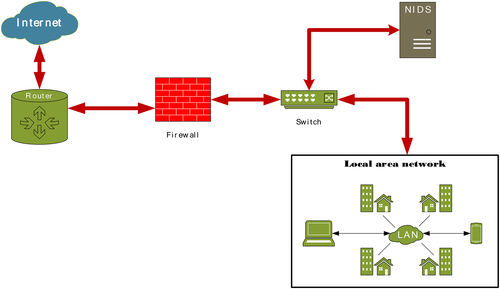
Description automatically generated

<https://www.youtube.com/watch?v=wUMObYAhQ4I>

A screenshot of a white box

Description automatically generatedA white text on a white background

Description automatically generated



<https://onlinelibrary.wiley.com/doi/full/10.1111/exsy.13066?saml_referrer>