COLLABORATION

Perth Conference 2017

AISA

COLLABORATION

Perth Conference 2017

Jill Slay

Developing Network Forensic Mechanisms for the Botnet of Things

AISA

- Work carried out by Slay, Sitnikova, Moustafa and Koroniotis

- Introduction
- Background to IoT and our work
- Machine Learning
- IoT Security Review
- Botnets in the IOT
- Network Forensics in the IOT
- Network Forensics Architectures
- Evaluation Metrics
- Results
- Conclusion

Overview

- The IoT is a network of interconnected everyday objects called “things”, that have been augmented with a small measure of computing capabilities. Lately, the IoT has been affected by a variety of different botnet activities.
- As botnets have been the cause of serious security risks and financial damage over the years, existing network forensic techniques cannot identify and track current sophisticated methods of botnets.
- This is because commercial tools mainly depend on signature-based approaches that cannot discover new forms of botnet. In the literature, several studies have conducted the use of Machine Learning (ML) techniques in order to train and validate a model for defining such attacks, but they still produce high false alarm rates with the challenge of investigating the tracks of botnets.
- This presentation investigates the role of ML techniques for developing a network forensic mechanism based on network flow identifiers that can track suspicious activities of botnets.
- The experimental results using the UNSW-NB15 dataset revealed that ML techniques with flow identifiers can effectively and efficiently detect botnets’ attacks and their tracks.

The Internet of Things

- The concept thought to date back to the early 1980s.
  - a vending machine selling beverages at the Carnegie Mellon University was connected to the Internet, so that its inventory could be accessed online to determine if drinks were available

- Term covering a multitude of devices and technologies.
  - Can be viewed as a collection of devices with low processing power and network communication capabilities.

- Increasing number of IoT devices.
  - In 2017 reached 8,3 billion, are expected to reach 20,4 billion in 2020.

Recent Events

- Mirai
  - a botnet consisting of 100.000 infected “things” that in October 2016, attacked and took out a good portion of the Internet’s high-profile services such as Twitter and Netflix by doing a DDoS attack on Dyn (DNS provider) [36].
  - Since then, Mirai has slowly been divided into smaller botnets, and new botnets have risen, such as BrickerBot, which as its name implies “bricks” an IoT device (permanently disables it) and Hajime

Our work

- Different security controls have been used for defining botnet events, including network forensic techniques and tools and intrusion detection and prevention systems.
- The existing techniques and tools basically use the principle of expert system which is generating specific rules in a blacklist and matching the upcoming network traffic against those rules.
- In our study, we investigate and analyse the role of ML techniques to construct a network forensic technique based on network flow identifiers (i.e., source and destination IP address/ports and protocols).

Machine Learning Techniques

- Machine learning techniques, learn and validate given network data for classifying legitimate and anomalous observation and have been utilized for building network forensic techniques, but there are still two challenges to be addressed:
  - producing high false alarm rates and
  - defining the paths of attacks, in particular botnet events [6] [22] [23] [24] [25].
- Machine learning algorithms include clustering, classification, pattern recognition, correlation, statistical techniques [6] [24] [26].
  - In clustering techniques, network traffic is separated into groups, based on the similarity of the data, without the need to pre-define these groups
  - In classification mechanisms they learn and test patterns of network data associated with their class label [6].

Our contribution

- The main contribution of our current work is the use of four classification techniques, so-called, Decision Tree C4.5 (DT), Association Rule Mining (ARM), Artificial Neural Network (ANN) and Naïve Bayes (NB) for defining and investigating the origins of botnets.

- The four techniques are used to recognise attack vectors, while the origins of attacks are linked with their flow identifiers as a network forensic mechanism.

Background – IOT Security (M. M. Hossain and M. Fotouhi and R. Hasan [38])

- Overview of open security challenges, forensic issues and attack surfaces in the **IoT**.
- A systematic analysis of attack surfaces and forensic issues in the IoT.
- Attack surface multiplies in IoT due to volume of devices, complexity, heterogeneity, interoperability, mobility and distribution of devices.
- Security constraints:
  Limitations based on hardware: Computational, Energy Constraint, Memory Constrain, Exposed hardware
  Limitations based on software: Embedded software constraint, Dynamic security patching,
  Limitations based on network: Mobility, Scalability, Diversity of devices, Diversity of connection protocols, Multi-protocol Networking, Dynamic network topology

IOT Security - Eyal Ronen et al. [38].

- The researchers investigate a possible attack vector for Philips Hue smart lamps, and a way that these IoT devices, which primarily use Zigbee as a communication protocol for the lamps to communicate with each other and their controllers.
- The attack vector they proposed and tested, takes advantage of exploits found in such devices and proposed that, under certain conditions (number of devices and relative distance), a malware can spread through a city, “jumping” from device to device, or even permanently disable (“brick”) them.

IOT Security - Yin Minn Pa Pa, et al. [39],

- They made observations on the nature and type of IoT devices being actively scanned through the Internet and went on to propose an IoT specific Honeypot named IoTPOT. Their solution included an IoTBOX a collection of virtual IoT machines (Linux OS), which helps make the Honeypot appear as a legitimate device.
- They observed a number of malware in action, most interestingly, they observed a repeating pattern, where a single host would perform the intrusion and information gathering process and the actual infection would be handled by a different host.

- However in literature review , it is evident that the field of IoT still being developed and as such various issues with this new and growing field are being discovered daily.

Botnets in the IoT

In the literature, there are a variety of techniques that researchers utilize to understand Botnets and to study them has been observed.

Ashkan Rahimian, et al. [40] studied Citadel (botnet whose main function was to steal bank and other credentials, also some spyware capabilities). They employed several code analysis techniques to measure the similarities between Zeus and Citadel

- Botnet scanning and identification was the main focus of Amir Houmansadr, et al. [41], who introduced BotMosaic, a tool following an architecture similar to Client-Server (clients are sensors that scan routers). Among other things BotMosaic, performs non-distorting network flow watermarking, for detecting IRC Botnets, which is the non-altering (towards network traffic content) process of “Marking” traffic so that it can be identified at a later time period..
-
- Botnets are capable of launching a number of attacks, like Distributed Denial of Service attacks (DDoS), Keylogging, Phishing and Spamming, Identity theft and even other Bot proliferation [4].

Some research has been conducted in developing ways to detect and identify the presence of botnets in a network. One such way is the utilization of machine learning techniques on captured packets which are often grouped into network flows (Netflows), in an attempt to distinguish between legitimate user traffic and botnet traffic [5].

A botnet's malware gets delivered to vulnerable targets through what is known as a propagation mechanism. Most commonly there exist two types of propagation, passive and active.

Various studies [18][19][22][23][24][25][26] have employed Machine learning techniques, to distinguish between normal and botnet network traffic and designing network forensic techniques and tools. In their work, Roux et al [18], created an intrusion detection system for IoT which takes into account wireless transmission data through probes.

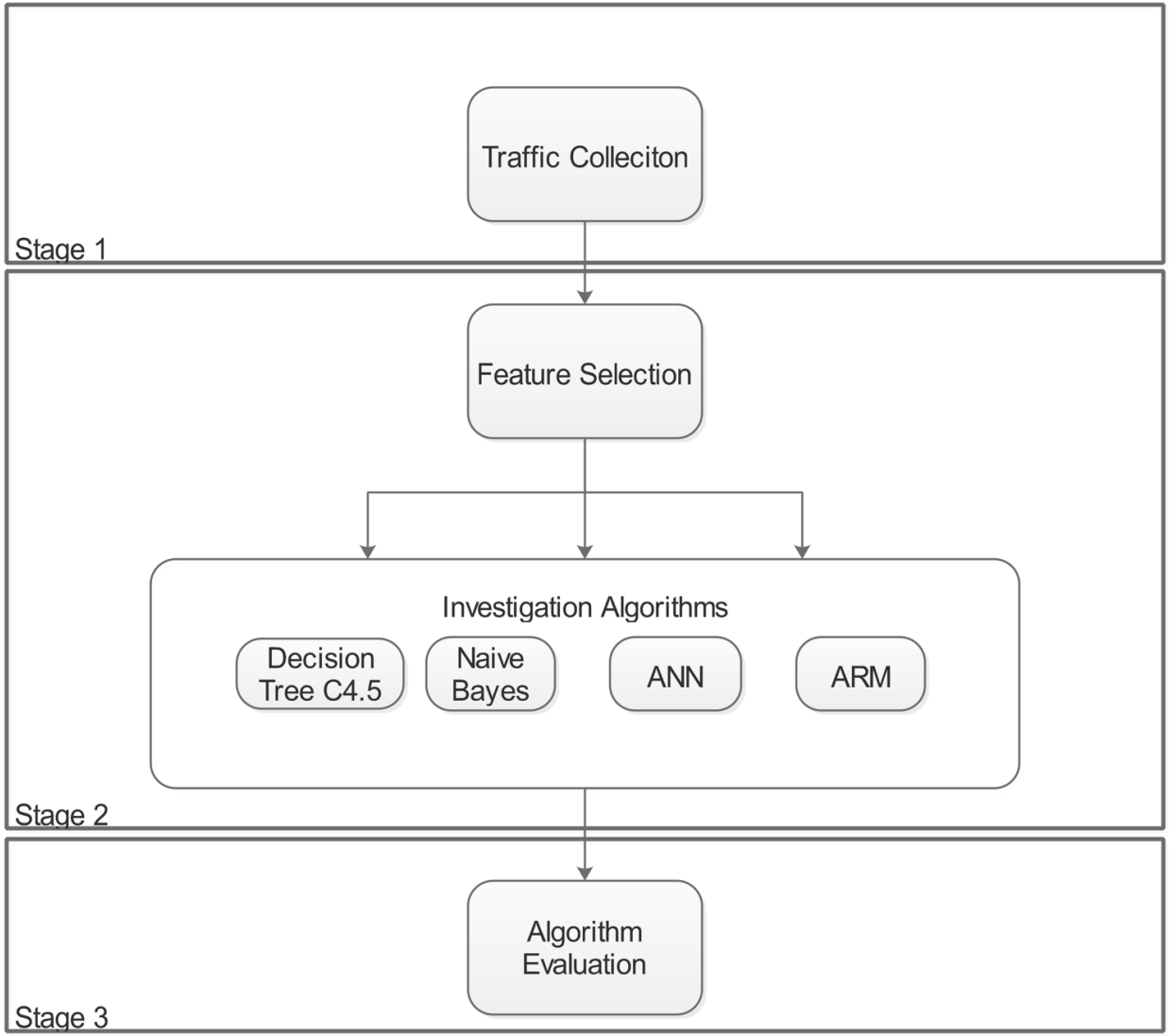
Lin et al [19], employed Artificial Fish Swarm Algorithm to produce the optimal feature set, which was then provided to a Support Vector Machine to be detect botnet traffic. They reported a slight increase in accuracy when compared with Genetic Algorithms for feature selection, but produced great improvement time-wise. Greensmith et al [20], proposed the utilization of Artificial Immune Systems as a way of securing the IoT.

Network Forensics in IoT

- Rana Khattak, et al. [27] tackled the problem of the ever-increasing size of network log files, by using parallel execution through Hadoop's MapReduce.
- S.Bansal et al. [28] proposed their own generic framework for detecting Botnets, focusing on Network Forensics techniques, such as packet capturing and inspection, which has the structure of a balanced framework, although it appears to be quite theoretical in nature.
- Saied et al [29] employed an Artificial Neural Network for developing a distributed topology of DDoS inspectors residing in different networks, to detect DDoS attacks, based on timing and header inspection.
- Divakaran et al [30] developed their own Framework for detecting such attacks (DDoS), by employing a regression model based on defined patterns. They did so, by grouping packets into network flows, and flows into sessions, based on timing of packet arrival, and then through regression, they identify anomalous patterns in traffic, also providing information on detection rates for different botnets.

Network forensic architecture and components

- The proposed network forensics mechanism includes four components: traffic collection network feature selection, machine learning techniques, and evaluation metrics, as below.



Traffic collection

- The immense volume of network traffic generated by todays networks, makes it necessary for a way to aggregate and summarize the captured packets, allowing for easier storage so that they can be used in construction of network forensic mechanisms.
- The network sniffing process needs to be conducted at key points of the network, particularly, at ingress routers, to make sure that the network flows that are gathered are relevant, which is determined by the source/destination IP address and protocol. This process helps, in investigating the origins of cyber-related incidents, and also lowers the necessary processing time.

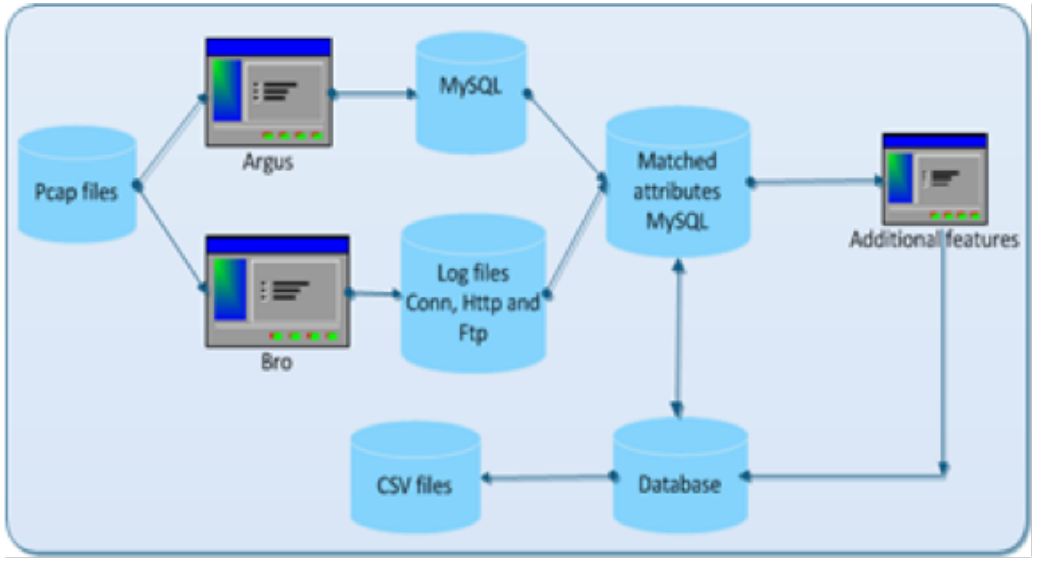


Fig.2: collecting network features of UNSW-NB15 dataset [15]

Network feature selection method

- Feature selection is the method of adopting important/relevant attributes in a given dataset. Method of feature selection is classified into filter, wrapper and hybrid of the first two.
- A filter feature selection mechanism denotes selecting suitable features without the use of class label, while a wrapper one depends on ML techniques [8][9]. An example of filtering methods for feature selection, is Information Gain (IG) and Chi-square (x^2).
- Wrapper methods study the performance of an algorithm that will be used in the end, as a criterion for selecting a suitable subset of the existing features.
- Intuitively, these methods split the features into subsets, and use these subsets to train the model that will be used after the pre-processing stage, using the error rate to score the subsets.
- Hybrid methods combine filter and wrapper methods to perform feature selection during the execution of machine learning algorithms.
- We use the information gain feature mechanism as it is one of the simplest methods that can adopt relevant features in large-scale data, as in network datasets. More precisely, Information Gain (IG) selects features, by calculates the apparent reduction in entropy, when the feature to be selected is used to split the dataset.

Machine learning techniques

For the classification stage, we use the Weka tool [42] for applying four well-known machine learning algorithms. These algorithms are briefly described as follows.

Association Rule Mining (ARM) [14]- is a classification algorithm, which is performed by generating rules of a form similar to $\{V_1,V_2,\ldots,V_n\} => \{C_1\}$, where $V_{1\text{-}n}$ are values of features and C_1 is a class value.

Artificial Neural Network (ANN) [11][12]- is a classification model which was based on the idea of the human neurons [11][12]. They usually are comprised of a number of neurons, which have multiple input flows and a single output, with some variants having multiple layers of units. The simplest form of an ANN is called a perceptron, taking the vector of attributes as input and classifying it.

Naïve Bayes (NB) [13]- classifies a record R_1 (collection of features) into a specific class C_2, if and only if the probability of that record to belong to that specific class, with respect to the record is greater than the probability of the record belonging to another class That is, $P(C_2/R_1) > P(C_n/R_1)$, with C_n being any class other than C_2.

Decision Tree C4.5 (DT) [11]- is a classification algorithm which produces a tree-like structure to determine the class chosen for a record. The attributes are used as the nodes of the tree and criteria are formulated leading from one node to the next, with the leaves being the Class value that is assigned to the record [11].

Evaluation metrics

- We utilize the confusion matrix [32] as a way of comparing the performance of the ML algorithms presented above
- An example of a confusion matrix is given in Table 1.

- In simple terms, it is a table which depicts the possible outcomes of a classification, which in our case is either '1', there was an attack detected or '0' normal network traffic, against the actual values of the class feature already present in the evaluation (testing) dataset.
- There are four condition that can be shown in a confusion matrix,
  - True Positive (TP), where the classifier has correctly identified the class feature and the value of that feature is positive (in our case there was an attack detected),
  - True Negative (TN), similar to TP but the value of the class feature is negative (normal traffic),
  - False Positive (FP), where the classifier identifies a record as an attack when, in actuality it is normal traffic and False Negative (FN), which incorrectly classifies an attack record as normal traffic

By combining the TP, TN, FP, FN values we are able to create two metrics, namely **Accuracy** and **False Alarm Rate**, which we can use to evaluate the Classifiers. These two metrics are calculated as follows:

•**Accuracy** represents the probability that a record is correctly identified, either as attack, or as normal traffic. The calculation of Accuracy (Overall Success Rate) is OSR= (TN+TP)/(TP+FP+TN+FN)

•**False Alarm Rate (FAR)** represents the probability that a record gets incorrectly classified. The calculation of the False Alarm Rate is FAR = FP+FN/(FP+FN+TP+TN)

	Actual Negative	Actual Positive
Predicted Negative	TN	FP
Predicted Positive	FN	TP

Table 1: Confusion matrix

Experimental results

Dataset used for evaluation and feature selection

- In order to compare the four aforementioned algorithms, we used the UNSW-NB15 dataset was designed at the Cyber Range Lab of the Australian Center of Cyber Security at UNSW Canberra [15]

- The dataset was produced by making use of the IXIA PerfectStorm tool, which produced a mixture of legitimate user network traffic and attack traffic, with the latter being categorised into 9 groups, Fuzzers, Analysis, Backdoor, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms.

- The dataset is comprised of 49 features, including the class feature, and the portion of it used, contains 257,673 (created by combining the training and testing datasets).

To test the classifiers, we performed Information Gain Ranking Filter (IG) for selecting the highest ten ranked features as listed in Table 2.

Table 2: UNSW-NB15 Features selected with Information Gain

Ranking	Feature selected	Feature description
0.654	sbytes	Source to destination transaction bytes
0.491	dbytes	Destination to source transaction bytes
0.477	smean	Mean packet size transmitted by source
0.464	sload	Source bits per second
0.454	ct_state_ttl	
0.444	sttl	Source to destination time to live value
0.439	dttl	Destination to source time to live value
0.429	rate	
0.409	dur	Record total duration
0.406	dmean	Mean packet size transmitted by destination

- The confusion matrices of the four classification algorithms are listed in Tables 3-6 on the training and testing sets of the UNSW-NB15 dataset.
- The Weka tool was used for applying the four techniques using the default parameters with a 10-fold cross validation in order to effectively measure their performance.

1. Our experiments show that Decision Tree C4,5 Classifier was the best at distinguishing between Botnet and normal network traffic. This algorithm makes use of Information Gain, to pick the feature which best splits the data based on the classification feature, during construction of the tree and at every node. The test showed that DT had the highest accuracy out of all the algorithms that were tested at 93.23%, and the lowest FAR at 6.77%.
2. ARM was the second-best classifier, having an accuracy of close to 86% and FAR just over twice that of the DT.
3. The Naïve Bayes classifier, which relies on probability to classify records in classes was third, with 20% less accuracy and close to 21% more false alarms than the DT.
4. Finally, the Artificial Neural Network was the least accurate out of the four algorithms that we tested, with accuracy and false alarm rate for this classifier showing a 30% differentiation from the C4,5 algorithm.

Normal	Attack	Prediction/Actual
31785	10894	Normal
12675	108654	Attack

Table 3: ARM confusion matrix

Normal	Attack	Prediction/Actual
84607	8393	Normal
9058	155615	Attack

Table 4: DT confusion matrix

Normal	Attack	Prediction/Actual
84101	8899	Normal
61380	103293	Attack

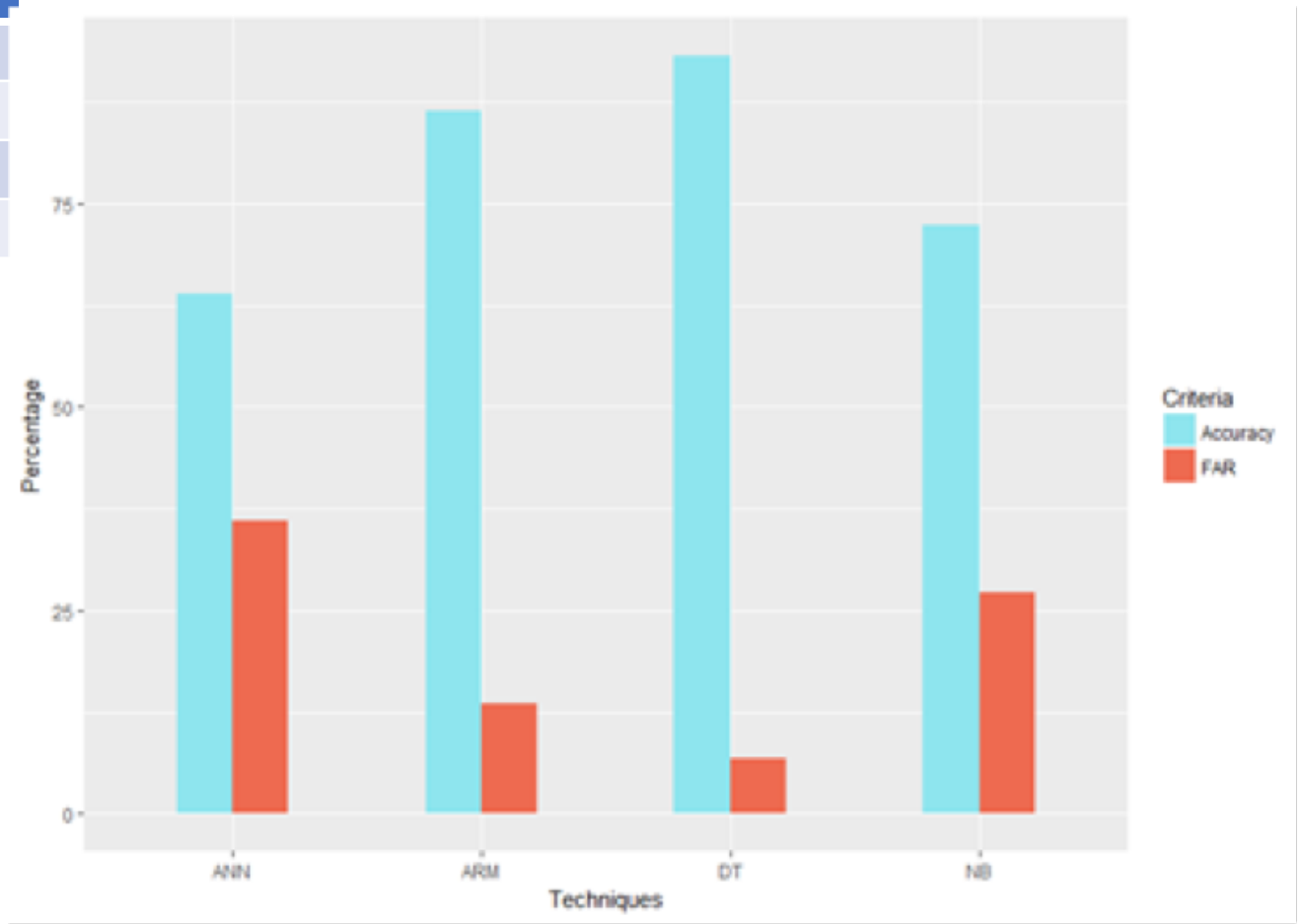
Table 5: NB confusion matrix

Normal	Attack	Prediction/Actual
2719	90281	Normal
2562	162111	Attack

Table 6: ANN confusion matrix

Table 7: Performance evaluation of four techniques

Classifier	Accuracy	False Alarm Rate
ARM	86.45%	13.55%
DT	93.23%	6.77 %
NB	72.73%	27.27%
ANN	63.97%	36.03%



- By combining the identifiers of a network flow, with the corresponding condition Label, the tracking of attack instances becomes possible.
- An example of the final form of the dataset is given in Table 8, which provides a number of flows and their classification label, taken from the UNSW-NB15 dataset.
- The network forensic technique that this paper illustrates, can assist network administrators, security experts or even law enforcement, to identify, track, report and even mitigate security incidents that threaten the security of their network.

srcip	sport	dstip	dsport	proto	Label
149.171.126.14	179	175.45.176.3	33159	tcp	0
149.171.126.18	1043	175.45.176.3	53	udp	0
175.45.176.3	46577	149.171.126.18	25	tcp	1
149.171.126.15	1043	175.45.176.3	53	udp	0
175.45.176.2	16415	149.171.126.16	445	tcp	1

Table 8: Flow identifiers associated with actual label for investigating attacks

Conclusion

- Here we discuss the role of machine learning techniques for identifying and investigating botnets.
- 4 ML techniques of DT, ANN, NB and ANN machine are evaluated on the USNW-NB15 dataset.
- The accuracy and false alarm rate of the techniques are assessed, and the results revealed the superiority of the DT compared with the others.
- The best machine learning techniques and flow identifiers of source/destination IP addresses and protocols can effectively and efficiently detect botnets and their origins as network forensic mechanism.