International Journal of
Advances in Scientific Research and Engineering (ijasre)

E-ISSN : 2454-8006

DOI: 10.31695/IJASRE.2019.33278

Volume 5, Issue 6
June - 2019

Implementation Hybrid (NIDS) System using Anomaly Holt-winter Algorithm and Signature based Scheme

[1]Mohamed Elmubarak, [2]Abdelrahman Karrar and [3]Nafeesa Hassan

[1]Nillian University, Khartoum, Sudan,

[2]Taibah University, Elmadinah,Saudi Arabia

[3]Nillian University, Khartoum, Sudan

ABSTRACT

An intrusion detection system is a security attack detection/prevention mechanism, it can be implemented into the software module or hardware module for the purpose of monitoring the systems or network for malicious activities. IDS can be categorized by monitoring resources to Host Intrusion Detection System (HIDS) and Network Intrusion Detection System (NIDS). HIDS are deployed to monitor local activities on the specific machine; on the other hand, NIDS placed into the central point on the network such as firewall to monitor network traffic. IDS also can be categorized depending on the detection method for anomaly and misuse (signature).

In this paper, we implement a hybrid implementation of IDS using Holt-winter anomaly algorithm and signature-based approach. Furthermore, a case study using Holt-winter anomaly based and signature misuse based schemes will be implemented and analyzed, finally, the result of the experiment will be shown.

Keywords: Intrusion Detection System, Network Intrusion Detection System.

1. INTRODUCTION

Intrusion Detection System (IDS) has been developed using various methods and techniques. The IDS analyses network traffic for each inbound/outbound packet and looks for abnormal activity. The advantage of this service is that it protects a user even if he or she is absent in the time of the attack. In addition, this service can generate alerts and notify users when discovering any malicious activities, allowing the system to recover the damage more quickly. IDS can be categorized by monitoring resources to Host Intrusion Detection System (HIDS) and Network Intrusion Detection System (NIDS). HIDS are deployed to monitor local activities on specific machine; on the other hand NIDS placed into central point on the network such as firewall to monitor network traffic. IDS also can be categorized depend on detection method for anomaly and misuse (signature). IDS can be used to identify the malicious or normal activity, and to generate alarms that notify system administrator in the case of intruder has been detected also it can be categorized based on architecture to centralized and distributed IDS, for the purpose of this paper we will focus on hybrid NIDS, we will show its components and structure. Furthermore, a case study using Holt-winter anomaly based and signature misuse based schemes will be implemented and analysed, finally the result of the study will be shown.

2. SIGNATURE BASED IDS

In the signature model the network traffic is scanned according to series of malicious packet sequence or bytes [1]. The main advantage of signature technique is the simplicity of implementation if we have a clear model for the network behavior we trying to protect. For example it can be use the signature to search for a particular pattern within exploit particular buffer overflow vulnerability.

When the signature based IDS find any matching in malicious signatures with the scanned activities it rises the alerts. The matching process can be done more efficiently using the modern computers, so the cost of the power required to perform this matching can be very low. For example, if the system we need to protect only use DNS, ICMP and SMTP, then the other signatures can be ignored during matching [1].

2.1 Snort: A Signature Based IDS

Snort is one of small, lightweight Intrusion Detection Systems which is written by Martin Roesch [2]. It is an open source Intrusion Detection System, also it can be configure as intrusion prevention and monitoring the networks in order to prevent attacks on these networks. It is a cross-platform network intrusion detection tool that can be deployed to monitor small TCP/IP networks and when detect suspicious network traffic Snort sends a real-time alert to syslog. It performs protocol analysis and content matching to detect intrusions [3].

2.1.1 Snort Components

Figure 2.1 shows the components of snort .In order to detect various attacks and produce output all of these components work together. They reside on top of the Libpcap promiscuous packet capturing library, which provides portable packet sniffing and filtering capability [3].

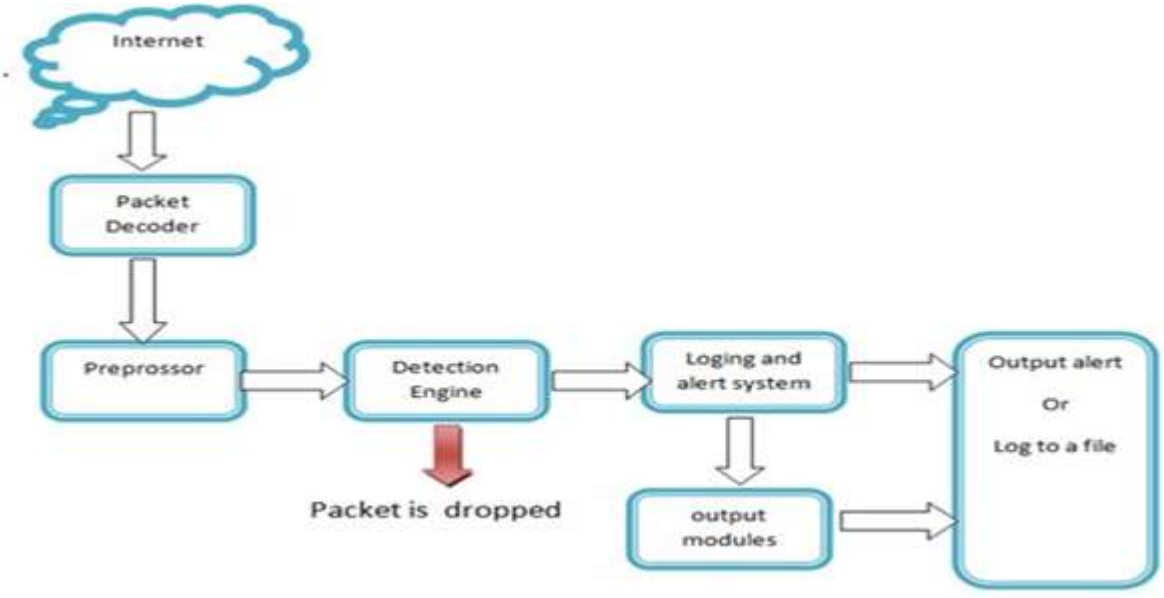

Figure 2.1: the component of snort [2]

According to [2][3] snort components are :

I. Packet Decoder

The main task of packet decoder is to take packets that come from various types of network interfaces and it prepares the packets to be preprocessed .The interfaces may be Ethernet, SLIP, and PPP.

II. Preprocessor

After the packets are passed from packet decoder to the second component which is preprocessor the preprocessor in turn arrange and modify the data packet according to the rules and passes it to another component which is a detection engine to find out if the packet is being used by an attacker.

III. Detection Engine

The most essential part of snort is the detection engine. If packets contain any malicious activity it will employ snort rules and the appropriate action will be taken; otherwise the packet is dropped, here appropriate action may be producing alerts [4].

The performance of detection engine can be effected by the following factors:

- Number of rules
- Cost of required Power of the device on which Snort is launched
- Internal bus speed used in Snort device.

DOI: 10.31695/IJASRE.2019.33278

- Network load.

IV. Logging and Alert System

In this component the packet may be used either to log activity or produce an alert this depend on what the detection engine detected in a packet. The logs reside in text files for example tcpdump style file.

V. Output module

Output modules is used save the alerts result. Depending on how you want to save output it can do different operations generated by the logging and alerting system of Snort. These modules can control the type of output generated by the logging and alerting system.

3. ANOMALY BASED IDS

The anomaly model is based on defining the network normal behaviour into a profile. The network traffic is classified in accordance with the predefined normal profile, then it is accepted or else it triggers the event in the anomaly detection. The accepted network behaviour is prepared or learned by the specifications of the network administrators.
The important phase in defining the network behavior is the IDS engine capability to cut through the various protocols at all levels. The Engine must be able to process the protocols and understand its goal. Though this protocol analysis is computationally expensive, the benefits it generates like increasing the rule set helps in less false positive alarms [4].
Profile generating is a process of creating network normal behavior and stores it into appropriate format. The profile is generated from network log that contain normal traffic, this process is done in training mode.

3.1 Holt-winter Algorithm

Work in [5] describes the holt-winter algorithm that used to generate the profile file that will be used in the anomaly implementation in the experiment of this paper. Holt-winter is an algorithm to generate the system profile. Holt-winter algorithm predicted values into time series; it uses a process known as exponential smoothing. All data values in a series contribute to the calculation of the prediction model. The Holt-Winters model uses a modified form of exponential smoothing. The algorithm perform three operations of exponential smoothing formulae to the series [8]. Firstly, the level (or mean) is smoothed to give a local average value for the series. Secondly, the trend is smoothed and lastly each seasonal sub-series (i.e. all the January values, all the February values. for monthly data) are smoothed separately to give a seasonal estimate for each of the seasons [5].
The exponential smoothing formulae applied to a series with a trend and constant seasonal component using the Holt-Winters additive technique are [8]:

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}\ s_t = \gamma(Y_t - a_t) + (1-\gamma)s_{t-p}$$

Where: a, β and γ are the smoothing parameters a_t is the smoothed level at time t b_t is the change in the trend at time t
s_t parameter represent the seasonal smooth at time t p represent the number of seasons per year
The forecast is obtained using the latest estimates from the appropriate exponential smoothies that have been applied to the series. So the forecast for time period $T + \tau$:

$$\hat{Y}_{T+\tau} = a_T + \tau b_T + s_T$$

Where:
a_T represent the smoothed estimate of the level at time T

b_T represent the smoothed estimate of the change in the trend value at time T s_T is the smoothed estimate of the appropriate seasonal component at T

4. Hybrid NIDS

The hybrid model consist of the anomaly and misuse detection methods combined together .The anomaly method is based on defining the normal network behavior which is predefined behavior into profile, then it is accepted or else it triggers the event as anomaly if not matched the excepted behavior.

DOI: 10.31695/IJASRE.2019.33278

The major drawback of anomaly detection approach is defining its rule set. The efficiency of the detection is depends the level of training that contain all potential behavior of legitimate traffic also the sized of generated normal predicted values into the profile. Thus correct detection depend on the detailed knowledge about the accepted network behavior and it need to be developed carefully by the system administrators. However anomaly works well when the rules are defined in precise way.in the other hand legitimate traffic can be changed and in this case anomaly detection will produce a large number of false alerts since new legitimate behavior is no defined on the profile yet.
Misuse detection is based on define a set of signatures or predefined knowledge about the attacks which is represented into database of signatures and rules which can be used to decide that a given pattern is belong to an intruder or not. Accordingly, misuse based systems are capable of getting the high levels of accuracy and minimal number of false positives in identifying even very subtle intrusions. Any tiny variation in known attacks may also affect the analysis if a detection system is not properly configured. Therefore we found that misuse intrusion detection system unfortunately cannot able to detect novel attack however this can be detected when used anomaly detection.
As a result of the presence of features in anomaly intrusion detection and lack thereof in misuse intrusion detection as well as via verse .The idea of the merger between the two methods lead to better performance and reduce the false alert [6].

5. Hybrid NIDS Implementation

Hybrid schema is consists of anomaly module and misuse detection module which they work together in examining the incoming packets in order to detect any potential malicious or attacks, figure 5.1 describes the schematic diagram of this combination.

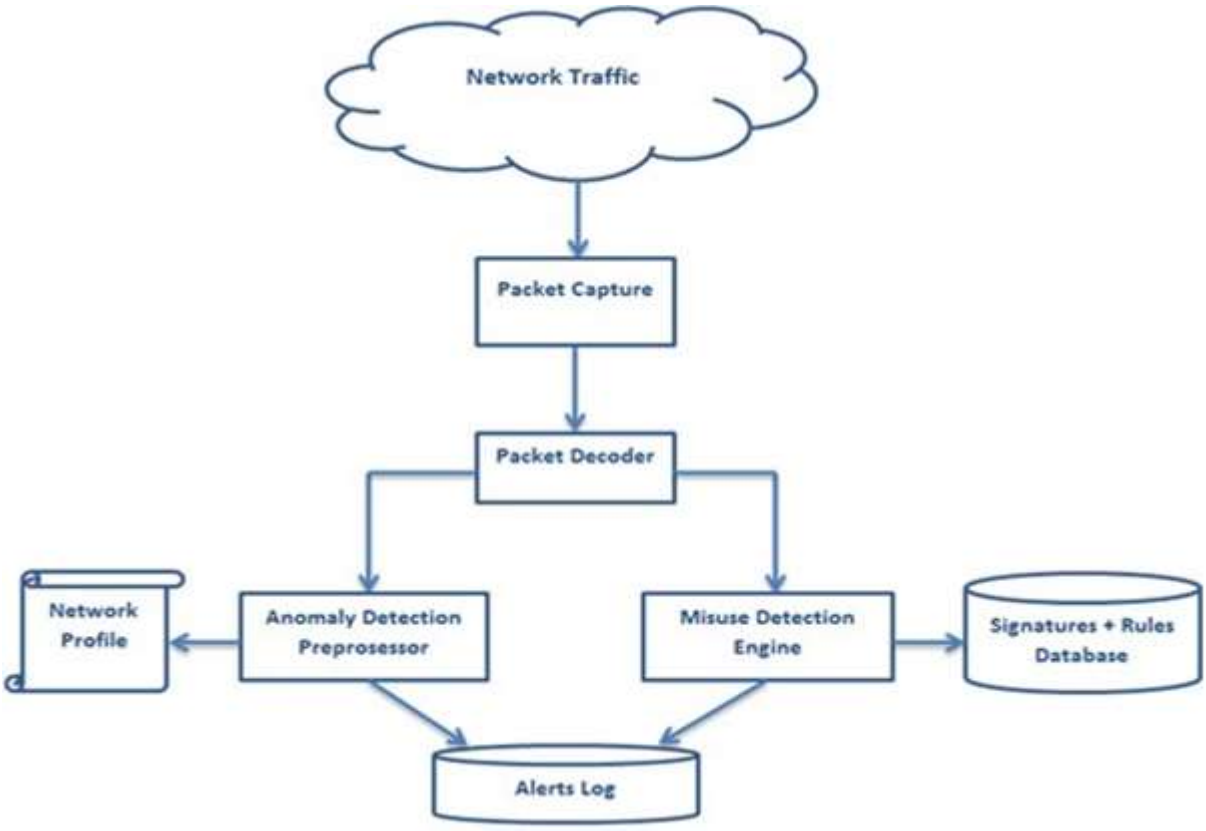

Figure 5.1: Hybrid model component and structure.

The system captures the packets from the network traffic using the packet capture, packet decoder is responsible of reassembling packet fragments and converting it into format that can be understood by anomaly and misuse detection modules. Anomaly module is act as pre-processor that uses the network profile to discover any deviation from the normal behaviour and generates alerts. On the other hand, the misuse module have a detection engine that uses database signatures to detect matched attacks into the incoming network traffic.

DOI: 10.31695/IJASRE.2019.33278

6. EXPERIMENT AND RESULTS

For this experiment, the virtual machines are deployed with Intel Core2 2.0 GHz processor,2 GB memory, running over Ubuntu Linux 12.04 operating system. The Snort tool has been used as well, where a different configuration represents a different model (anomaly, misuse and hybrid).

The three models utilized the DARAPA dataset with normal and attack traffic. The data of Thursday, week 3, containing only normal traffic, is inserted into each model to generate false positive and true positive alerts. Then, the data of Thursday, week 5, containing only attack traffic, is inserted into each model to obtain the false negatives and true negatives of each.

The day was divided into intervals of minutes and measured from 4 am - 23 pm, giving a total of 1140 minutes. Any traffic measured during this time on Thursday of week 3 will be considered as normal, so that any model that generates alerts in these intervals will be taken as a false positive alert.

Likewise, the same number of intervals will be considered as attacks in the Thursday of week 5, so that any model that does generate an alert during those intervals will be taken as a false negative alert.

6.1 Anomaly Model Results

The experiment carried out over the anomaly detection model was done in two phases:

Phase one, the positive alert is tested by reading the log traffic on Thursday in week3. It then generated 887 true positive alerts and 253 false positive alerts for a total of 1140 intervals.

Phase two, the negative alert is tested by reading the log traffic on Thursday in week5.

Accordingly, it generated 764 true negative alerts and 367 false negative alerts for the total 1140 attack intervals. The anomaly detection rate was 53,764 packets per second.

6.2 Misuse Model Results

In the same way, the experiment of the misuse model was also divided into two phases:

In phase one, the misuse model is tested against positive alerts entering the traffic of Thursday in week three. Accordingly, it generated 1140 true positive alerts and zero false positive alerts for a total of 1140 intervals.

In phase two, the negative alert is tested by reading the traffic of Thursday at week five. It then generated 395 true negative alerts and 745 false negative for a total of 1140 intervals. The misuse detection rate was 103943 packets per second

6.3 Hybrid Model Results

A final experiment is performed over the hybrid model, giving the following results:

In phase one, a test for positive alerts was produced test where the traffic of Thursday in

Week 3 was used. Accordingly, this generated 887 true positive alerts and 253 false positive for a total of 1140 intervals.

In phase two, the false negative is tested by reading the traffic on Thursday at week 5. It then generated 806 true negative alerts and 334 false negative alerts for a total of 1140 intervals.

The hybrid model detection rate was 49642 packets per second. Table 5.1 shows the comparison between the three models based on detection rate and false alerts.

	Anomaly Detection	Misuse Detection	Hybrid Detection
True Positive Rate	77.81%	100%	77.81%
False Positive Rate	20.61%	0%	20.61%
True Negative Rate	67.02%	31%	70.70%
False Negative Rate	32.19%	65%	29.30%
Detection Rate	53764/sec	103943/sec	49642/sec

Table 6.1: Comparing between three detection models in term of false alarms and detection rate.

DOI: 10.31695/IJASRE.2019.33278

6.4 Result Discussion

The experiments were carried out over the three models with a focus on three factors, which are false positive alerts, negative alerts and detection rate for each model.

As seen in Figure 6.1, the anomaly detection model produced a large number of false positive in comparison to the misuse detection model. This is due to the fact that the anomaly model counts the number of packets in the given time interval. It then compared this with the maximum and minimum values of the predefined data in the network profile, and so any minor change in user behavior can cause a higher number of false positive alarms.

The second phase of the test over the anomaly detection model is performed only on attack traffic to examine false negative alerts. This generated a large number of alerts. Anomaly detection compares the counts of traffic within the time interval, so that any deviation from the maximum or minimum is considered an attack thus generating alerts. However, the drawback of this approach is that if the attack does not exceed the maximum or minimum of the profile, the attack will not be detected and will be considered as a false negative alert. On the other hand, this model is advantageous in its ability to detect zero-day attacks.

As we shown in Figure 6.1 the misuse detection model does not produced any false positive alert after phase one of the test was done, this due to fact that misuse detection model based on modelling the attack by signatures or rules using unique characteristics and since week three of DARAPA dataset contain only normal traffic so misuse detection does not provide any false positive alerts. After implementing phase two of the test misuse generates

a numbers of alerts after matching the defined rules, the advantages of this model is that it does not produced any false positive alerts and the disadvantages of this model is it can only detected known attacks which are well defined into rules or signatures.

Hybrid model implement both of anomaly and misuse detection models together so after implementing phase one of the test it produced the identical number to anomaly model in false positive alerts, in the phase two of experiments the number of discovered attack was more than both of two models, and the reason for that is anomaly can detect unknown attacks and the misuse model detect predefined attack, so the hybrid model take advantages of both models and detect more numbers of the attacks than each model alone, but the advantages of the hybrid model is it generates much numbers of false positive than the misuse model alone Figure 6.1 shows these result.

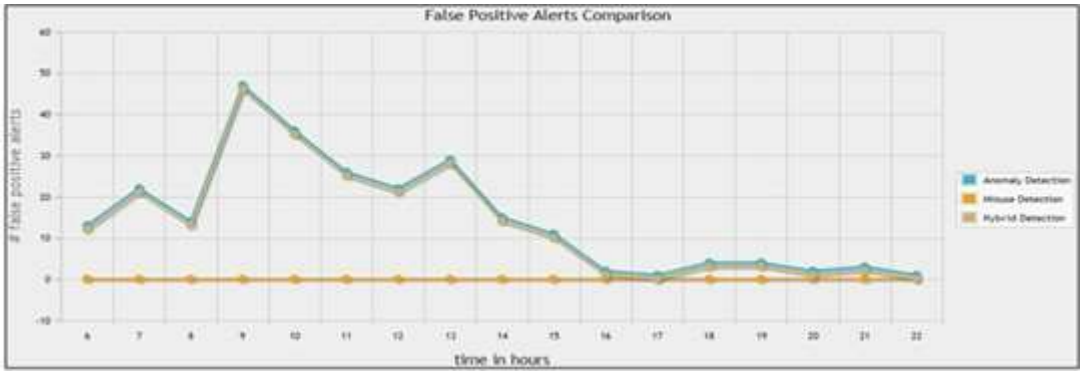

Figure 6.1: False positive comparison between anomaly, misuse and hybrid detection models.

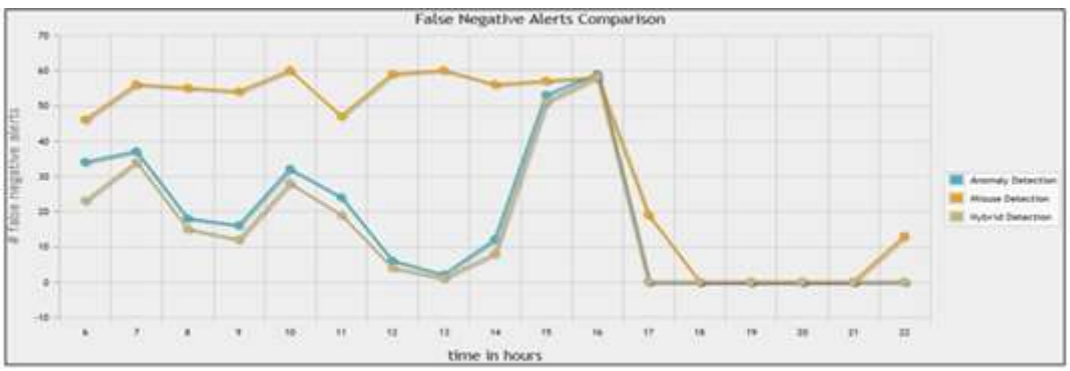

Figure 6.2: False negative comparison between anomaly, misuse and hybrid detection models.

DOI: 10.31695/IJASRE.2019.33278

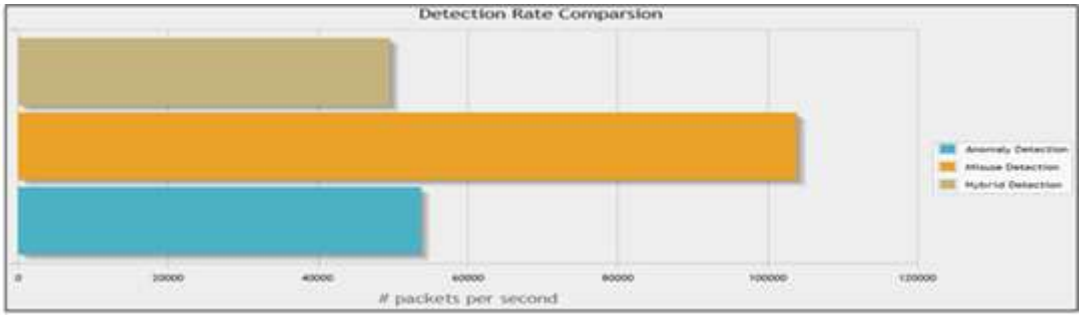

Figure 6.3: Detection rate comparison between anomaly, misuse and hybrid detection models.

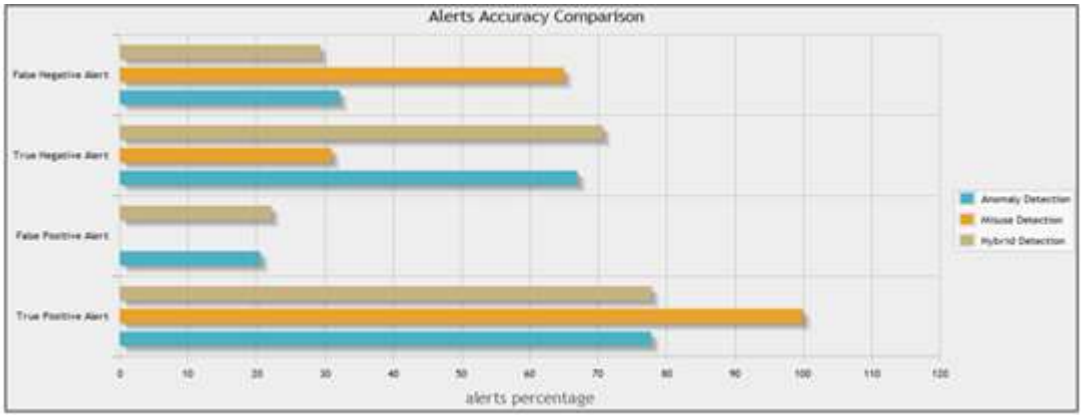

Figure 6.4: Alerts rate comparisons for anomaly, misuse and hybrid detection models.

7. CONCLUSION

In this paper we implement a hybrid implementation of IDS using holt-winter anomaly algorithm and signature based approach. Furthermore, a case study using Holt-winter anomaly based and signature misuse based schemes is implemented and analyzed. The hybrid model combine both of anomaly and misuse be implementing the model through snort IDS using Holt-winter algorithm and signature based scheme. The results showed that the hybrid scheme was more efficient in term of false alerts and detection rate.

7.1 FUTURE WORK AND RECOMMENDATIONS

7.1.1 It is known that, unlike the anomaly approach, the misuse approach cannot detect an unknown, novel attack. Since both detection methods could be work together in a single hybrid model, it is recommended that any attack alerts detected by the anomaly detection module is passed to an additional signature-generating module. A signature can then be assigned to that attack and added to the misuse signature database, so that the attack could later be recognized by the misuse detection module.

7.1.2 The experiment has been performed over DARAPA dataset which is offline traffic, the hybrid model can be tested on the live network traffic in the future.

REFERENCES

1. Dubey, S. and Tripathi, N., 2015. A Survey on Intrusion Detection Systems. International Journal of Scientific Research in Science, Engineering and Technology,(1), pp.29-40.

2. Martin Roesch, Snort -Lightweight Intrusion Detection for Networks, 13th USENIX Systems Administration Conference – LISA '99, Seattle, Washington, November 1999.

3. Mehra, P., 2012. A brief study and comparison of snort and bro open source network intrusion detection systems. International Journal of Advanced Research in Computer and Communication Engineering, 1(6), pp.383-386.

DOI: 10.31695/IJASRE.2019.33278

4. Rehman, R.U., 2003. Intrusion detection systems with Snort: advanced IDS techniques using Snort, Apache, MySQL, PHP, and ACID. Prentice Hall Professional.

5. Stallings, W., Brown, L., Bauer, M.D. and Bhattacharjee, A.K., 2012. Computer security: principles and practice (pp. 978-0). Upper Saddle River (NJ: Pearson Education).

6. Szmit, M., Adamus, S., Szmit, A. and Bugała, S., 2012, September. Implementation of Brutlag's algorithm in Anomaly Detection 3.0. In 2012 Federated Conference on Computer Science and Information Systems (FedCSIS) (pp. 685-691). IEEE.

7. Agarwal, N. and Hussain, S.Z., 2018. A closer look at Intrusion Detection System for web applications. Security and Communication Networks, 2018.

8. Paraschiv, D., Tudor, C. and Petrariu, R., 2015. The textile industry and sustainable development: a Holt–Winters forecasting investigation for the Eastern European area. Sustainability, 7(2), pp.1280-1291.

DOI: 10.31695/IJASRE.2019.33278