Hybrid Honeypot Malware Detection System Based On Machine Learning

Neha Titarmare, Diksha Bhagat, Neha Chaudhari, Samruddhi Sarode, Pranay Tagde, Prasad Lonare,Vishal Binekar
Computer Science & Engineering Department, Rajiv Gandhi College of Engineering and Research, Nagpur,Maharashtra, India.

ABSTRACT

The latest wireless technology is expanding smart phone technology and burgeoning mobile cloud technology. Mobile cloud computing provides a lot of advantages in the future, but it's also easy for hackers to take full control of a lot of other users. Data privacy is important. While data security is intended to be secure, the biggest disadvantage for consumers is that when the computer is connected to the internet, an intruder can easily steal data from the desired target. So, in order to provide greater security, a mix of Hybrid Intrusion Detection System (HyInt) and Honey pot networks has been introduced into the Mobile Cloud Environment with the primary goal of mitigating both unknown and known assaults.The study work's execution provides a pure perspective on the algorithm's security and quality products that was not included in the prior research work. Intensive statistical analysis was carried out as part of the research to demonstrate the consistency of the suggested algorithm. The implementation and evaluation results show that there is plenty of room for more study in the cloud-based Intrusion Detection System. The created algorithm can be utilized to effectively monitor the network's activity in a high-security cloud environment intended for army and banking purposes.

Keywords – Hybrid Intrusion Detection System, Honeypot Networks, Signature and Anomaly based detection, Mobile Cloud Computing, Performance.

1. INTRODUCTION

A honeypot can be used in network security to uncover new attacks that Intrusion Detection Systems or network firewalls may not be able to detect using the previous static protection rule method. When designing IDS (Intrusion Detection Systems) and Firewalls, it is critical to consider the corporate defensive rules for going past the honeypot. Computer networks are prone to a variety of exploits that might render them insecure or prevent them from performing their intended function. Intruders and attackers have become increasingly agitated about network security and obstacles.Enterprises, organisations, and, more importantly, finance departments, have an essay solution to adopt various hardware and software for network security providers such as firewalls, variants of intrusion detectors, and [17] Virtual Private Networks to have a better and improved security. These systems, on the other hand, work nonstop to keep confidential information out of the hands of unscrupulous intruders, and to warn of new attacks as they occur.

Mobile Virtualization is the most advanced feature emerging in today's world, and its applications for smartphones are growing by the day. The number of mobile users is growing all the time since it makes work easier and faster, as well as providing the latest technology that is continually evolving

and allowing users to access all of their apps via the network from anywhere in the world.Mobile cloud computing has a major advantage in that it is very versatile, and we can access data and share information anywhere in the world even if we are not connected to the internet. It also offers cost-effectiveness in that use and maintenance are comparatively low, and real-time data availability, where all user information is available in real-time on our mobile device when connected to the network, from which we can update and acquaint ourselves.Despite all of the hoopla around the MCC, it has a serious flaw in terms of privacy and security, which contributes to trust issues for consumers and organizations. As the globe evolves, so do the number of hackers.

Similarly, companies are introducing new items and techniques for protection in the cloud computing environment, where cloud computing services are provided on a pay-as-you-go basis. When used together, how can the Hybrid Intrusion Detection System (HyINT) and Multi-Honeypot Network (MHN) improves security in Mobile Cloud Computing? Honeypot networks are used to achieve stronger defense in depth protection and overall security of the cloud environment. The analysis of attack tactics is identified in the honeypot networks network as necessary for countermeasures. Many destructive threats, such as DDOS, XSS injection, and SQL injection, can't be completely avoided, but they can be minimized.There are numerous ways to secure it from hackers, but IDS is the most important and widely used method for detecting bad code in a network, and it plays a significant role in protecting the cloud environment from attackers.

2. LITERATURE REVIEW

[1] "DDOS Detection Using Machine Learning Technique," by Sagar Pande. Numerous attacks on network infrastructures are identified in this problem. Attacks on network availability, confidentiality, and integrity are among them. A distributed denial-of-service (DDoS) attack is a persistent attack that degrades network availability. The problem is approached using the Command and Control (C&C) method, which is utilized to carry out such an attack. To detect these assaults, many researchers have developed various ways based on machine learning techniques. In this paper, a DDoS attack was carried out with the ping of death approach and detected with the WEKA tool utilizing machine learning techniques. The NSL-KDD dataset was used in this experiment as a result.The normal and assault samples were classified using the random forest technique. The samples were accurately categorized in 99.76 percent of the time.

[2] "A Survey of DDOS Attacks Using Machine Learning Techniques," Arshi M1, et al. DDOS stands for Distributed Denial of Service assaults in this problem. These assaults are becoming more complicated, and the number of them is predicted to grow every day, making it difficult to detect and counteract them. A sophisticated intrusion detection system (IDS) is required to identify and recognize aberrant internet traffic behavior as the solution to the problem. The procedure is supported in this article using the most recent dataset covering the most common types of DDoS attacks, including (HTTP flood, SIDDoS). This research incorporates well-known grouping techniques such as Nave Bayes, Multilayer Perception (MLP), and SVM, as well as decision trees.The end result is In the future, a full analysis of data sets covering the most recent forms of assaults, such as HTTP flooding, SIDDoS, Smurf and UDP flooding, will be conducted using deep

learning techniques on data acquired from the college network.

[3] "A DDoS Attack Detection Method Based on Machine Learning," Jiangtao Pei. The concern identified in this study is that, with the rapid advancement of computer and communication technology, the harm caused by DDoS attacks is becoming increasingly serious. The problem is handled by proposing a machine learning-based DDoS assault detection solution that incorporates two steps: feature extraction and model detection. The DDoS attack traffic characteristics with a substantial proportion are recovered in the feature extraction stage by comparing the data packages sorted according to rules. The experimental findings reveal that the proposed machine learning-based DDoS assault detection approach has a good detection rate for the current popular DDoS attack.

[4]Swathi Sambangi's paper "A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression" identifies the problem of recognizing Distributed Denial of Service (DDos) assaults as a classification problem in machine learning. The task of detecting DDoS attacks, which is relevant to Cloud Computing, is a substantial challenge due to the computational complexity that must be addressed.The research objective is to investigate the problem of DDoS attack detection in a cloud environment by using the most popular CICIDS 2017 benchmark dataset and applying multiple regression analysis to build a machine learning model to predict DDoS and Bot attacks using a Friday afternoon traffic logfile. As a result, a prediction accuracy of 97.86 percent was reached using this ensemble model for the Friday morning dataset. The prediction accuracy for the Friday afternoon log file is 73.79 percent for 16 attributes derived by information gain-based feature selection and regression analysis-based ML model.

[5] "Detection of DDOS Attacks Using Machine Learning Methods," Tuba Aytaç. The difficulty in this study is to figure out what the problem is. People's need to communicate with one another helps to technological advancement, and it has made the internet notion a vital part of our daily lives. Cyber-attacks from extranets to enterprise networks or intranets used for personal purposes can result in monetary loss as well as intangible damage. It is vital to take all necessary safeguards in order to minimize losses by detecting attacks early. This study's approach to the problem is to look at the success rate of intrusion detection systems utilizing various methodologies. The CICDDoS2019 data set was used in this study, and DDOS attacks in that data set were compared.Artificial Neural Networks (ANN), Support Vector Machine (SVM), Gaussian Naive Bayes, Multinomial Naive Bayes, Bernoulli Naive Bayes, Logistic Regression, K-nearest neighbor (KNN), Decision Tree (entropy-gini), and Random Forest methods were used to investigate threat detection success rates. The Fwd Pkt Len Std, Fwd Seg Size Min, Bwd Pkt Len Std, Bwd Pkt Len Min, CWE Flag Count, Bwd Seg Size Avg, Bwd Seg Size Min properties were found to be the most distinct values for determining DDOS attacks after analysing five data sets that had the highest accuracy rate in data that had been trained with different algorithms.

[6] "A Study on the Evolution of Ransomware Detection Using Machine Learning and Deep Learning Techniques," Damien Warren Fernando et al. The destructive nature of ransomware, the difficulty of reversing a ransomware infection, and the importance of detecting it before it infects a system are the key motives for this investigation in this paper. Because machine learning and deep learning can detect zero-day threats, the strategy for the problem is to look into machine learning and deep learning methodologies when it comes to detecting ransomware. These methods can be used

to create prediction models that can learn how ransomware behaves and use that knowledge to find variants and families that haven't been seen before.In this study, we look at some of the most well-known research studies that all use machine learning or deep learning to detect ransomware virus. As a result, we conducted tests to see how malware evolution affects the research topics described. We also discussed ransomware's new paths and how we expect it to expand in the coming years, such as its extension into IoT (Internet of Things), with IoT becoming more connected into infrastructures and into people's homes.

[7] "Ransomware Detection Using Dynamic Analysis and Machine Learning: A Survey and Research Directions," Umara Urooj, et al. A Survey and Research Direction Ransomware is a notorious piece of software that has gained notoriety due to its fatal and irreversible consequences on its victims. The irreversible damage inflicted by ransomware necessitates early detection of these attacks. The problem is tackled in a unique way. This research presents information on the datasets collected from various sources and used in ransomware detection tests on various platforms.This research is additionally unique in that it provides a survey of ransomware detection studies that use machine learning, deep learning, or a combination of the two techniques while using the benefits of dynamic analysis for ransomware detection. The research presented here takes into account ransomware detection investigations completed between 2019 and 2021. This research gives a comprehensive list of future directions, paving the path for future research. As a result, this study aims to produce a user manual that can serve as a guide for academics looking to work with existing technologies in the field of ransomware attack detection. It can assist them in designing more effective ransomware detection algorithms while taking into account existing solutions.We will continue to investigate the importance and contribution of static analysis in the detection of ransomware threats using machine and deep learning technologies in the future.

[8] "Ransomware: Recent Advances, Analysis, Challenges, and Future Research Directions," by Craig Beaman et al. The COVID-19 epidemic has resulted in a massive increase in the amount of ransomware attacks, according to this paper's problem. Various institutions have been targeted, including healthcare, finance, and government. There could be a variety of causes for the rapid increase in attacks, but it appears that working remotely in home-based environments is one of them. Recent developments in ransomware analysis, detection, and prevention were investigated as a solution to the problem. Honeypots, network traffic analysis, and machine learning-based approaches were discovered to be the main focus of state-of-the-art ransomware detection strategies.It was also discovered throughout the studies that ransomware may be simply developed and used. Finally, we discussed current research issues as well as future research goals in the realm of ransomware.

3. RELATED WORK

We present background and work related to the suggested solution in this study in this part. This is where we talk about Mobile Cloud Computing and its security concerns. 2.1 Mobile Cloud Computing Takes Center Stage Mobile Cloud Computing is the current and trending technology all over the world, and it offers a number of advantages that are quite beneficial in terms of improving the user experience. [2] From which it derives specialised features such as storage, smartphone mobility via wireless or internet connection, and pay-as-you-go service. Similarly, Juniper Research reports on the expanding use of mobile computing, noting that demand for cloud based mobile apps in the public

and commercial sectors is likely to rise to 9.5 billion dollars by 2014, while it is expected to rise even higher in the near future. Similarly, smartphone applications have grown in number in recent years, with apps in a variety of categories including entertainment, social media, online streaming, banking, news, and so on. The main reason for this is that mobile computing is capable of providing the subscriber with a resource when and how they need it, based solely on user organization.

According to a 2009 study by International Data Corporation (IDC), 74 percent of IT administrators and Chief Information Officers (CIOs) believe that user privacy concerns are the key risks that have kept most companies from adopting virtualization. Technology, hardware, and communications are three essential aspects that assist mobile computing. Hardware includes gadgets such as smart phones and other portable devices that clients can use. Consumers are gradually adopting PDAs, thanks to the rapid advancement of the wireless network [3] According to the Allied Business Intelligence estimate, more than 2.4 billion customers would utilize a mobile device to access cloud computing platforms in 2015. Similarly, Google emphasizes some cloud based products for consumers and businesses, including a vital item for mobile phones known as Android OS, as well as numerous programmers such as Google MapsStreets, and others. Similarly, Google Stadia is a new technology that is a cloud-based gaming service that does not require any hard ware and only requires an internet connection to connect [4] the basic techniques utilized in the technology business, such as parallelization model, virtualization, and mass production, are the three primary strategies for cloud computing, as shown in Figure 1.

Importance of Cloud Security

Since individuals use mobile phones in cloud environments, they are vulnerable to a variety of external dangers. Information privacy and authentication should be understood by regular users and software developers, as if they are aware of the privacy implications, there will be no difficulties with hackers. People nowadays are unaware of how to use technology and the advanced functions available on their cellphones and PDAs. Mobile protection can be performed by a variety of security elements, including app installation, such as anti-virus software. [1] [5]

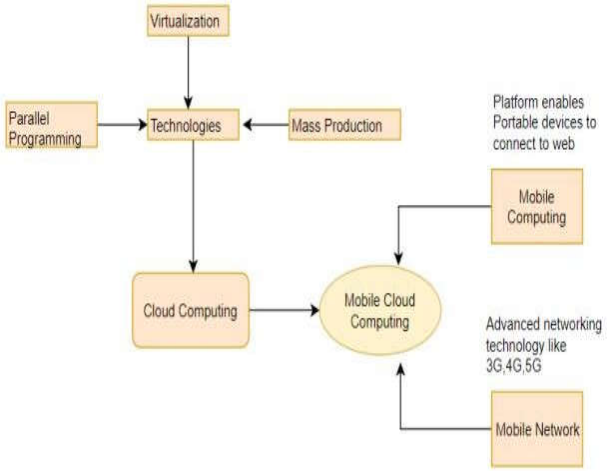

Figure 1: Architecture of Mobile Cloud Computing

Application security and data security frameworks are the two types of security frameworks for Mobile Cloud Computing (MCC).

Storing data in a database in a virtual environment without revealing any details is more difficult for mobile users. An authentication method is used to verify that when never a user transfers a file to a cloud server for her ring with many clients, it should also be verified that the user viewing the file is a trustworthy client. Scalability is a network's ability to interact with clients in a flawless manner [11].

Similarly, the most recent security technologies for online services, such as VPN usage, password encryption, authentication, and entry command, should be introduced to provide unitinterrupted services against various threats, such as DOS attacks and data theft [9] when such attacks occur, cloud services must provide a backup and restoration solution to restore client confidence. Figure 2 depicts recent security challenges and current approaches in the table below.

Security Issues		Current Approaches
Mobile Cloud	Platform Reliability	Authentication and access control, Privacy and data protection.
	Privacy and Data Protection	Key management and data encryption. Integrating the current security technologies.
Mobile Terminal and Network	Malware software	Detection and prevention CloudAV
	Software Vulnerabilities	Installing the system patches, checking software legitimacy and integrity.
	Information Leakage	Data Encryption and Security Protocol

Figure 2: List of Security Issues

Potential of Intrusion Detection

System in the Cloud An incursion is any attempt to breach the CIA of a device or network, and there are numerous types of intruder attacks. The most prevalent is (DOS) attacks, which prevent legitimate users from accessing internet based services. [6] In the virtual environment, an attacker can send many attempts to authenticate VMs using cyborgs, causing normal users to be unable to access them.Intrusion Detection and Prevention Systems (ID/PS) that are still available were unable to provide the required degree of protection and performance. Pandeeswari and Kumar (2016) used a Fuzzy Mean Clustering-based ANN to identify breaches in the cloud, where IDS is often implemented on end host cloud servers using the following methods. [5] [7]

Potential ransomware will be prevented from employing traditional HIDS based on signature matching methods by applying authentication procedures. Complex assessment based on existing IDS can be avoided by testing the controlled computer with the help of the security process.[8] Signature matching procedures necessitate proper monitoring, after which a second layer of protection (Modiand Patel, 2013) links new NIDS tools with older anomaly detection methods to detect cyber attacks in a network. Similarly, certain cloud protection services, such as Snort IDS, fail to distinguish VM attacks that target everyone from individual residents to distinct physical servers. The many forms of Cloud IDS are depicted in Figure 3[9].The effectiveness of an ID is enhanced by a hybrid intrusion detection system. Combining signature based techniques with anomaly based techniques can considerably improve this. The ability to withstand new, unknown attacks by leveraging existing knowledge gained from previous attacks. Even though it has the problems of all solutions that use centralized control in a dispersed setting [10], X.Wang et al proposed a methodology that is related to the central management approach.

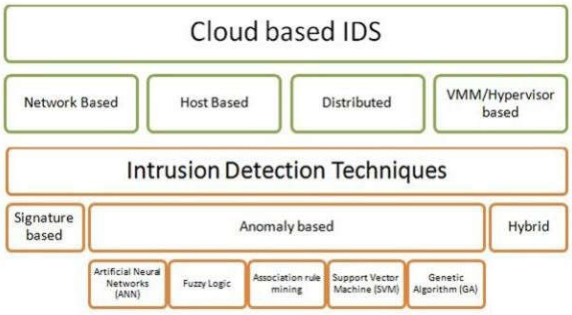

Figure 3: Structure of Intrusion Detection System in Cloud environment

Modi et al. started an approach for stepwise intrusion detection as well. Hybrid IDS is more useful in terms of vulnerability security and performance since it pre-processes packets and sends them to signature based IDS after comparing them to patterns that have already been detected. Previous systems were limited by the fact that they

couldn't be properly built to handle new forms of attacks, which is also a time consuming operation that takes too long to investigate suspected attempts.

Honeypot Uses of Intelligence

Honeypots are a complex idea in network security; such a device strives to gather information about infiltration attempts. Minimum interaction honeypots, which just simulate the communication layer, to heavy interaction honeypots, which run a full operating system, are available. One of the key reasons for employing a cloud service is together high and low communication honeypots used in a cloud environment to evaluate assaults, and they must check that the dispersed packets are legitimate after they are moved to HoneyCY as their transition to the cloud [12].Similarly, it is made up of three design layers: HoneySrv gathers honeypie devices and information, and HoneyV manalyzes malware collected Brown et al mentioned a number of virtualization systems used in honeypot sensors, while Saadi et al presented an IDSf or a smartphone device that included Honeycomb, HoneyNet, and HoneyD honeypots. [13] [14].

4. ANALYSIS OF PROPOSED MODEL

TYPES OF HONEYPOTS

A. LOW INTERACTION HONEY POTS

Low-interaction honey pots in an aggressive expansion are simple yet can save time due to intruder detection, and the interaction honey pot imitate can be reduced with particular commands. Honeyd is an example of a low-interaction [5] honey pot. Taking advantage of the low involvement honeypot allows invaders to imitate with services with limited interaction. Because of the minimal degree of contact and essentially system breach, the goal of this type of honeypot is to collect data from the first step of an attack. Data regarding the threat's motivation is rarely obtained.

An IP address is required for the virtual honeypot software process. To share a single run, multiple virtual honeypots often use multiple IP addresses and network interfaces. As a result, the virtual honeypot is installed on a single physical system and network address translation is performed via a firewall or other means. The majority of high-interaction honeypots can completely compromise the production system, whereas low-interaction honeypots can only simulate virtual systems due to their restricted capabilities.

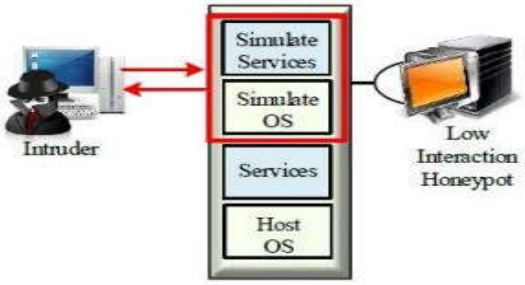

Figure 4: Low Interaction Honeypots

Honeyd's main task is to issue warnings, the majority of which are accurate and accurate attack alerts. Honeyds can detect any action on any User Datagram Protocol (UDP) or Transmission Control Protocol (TCP) port by default, and some of the activities are also written in ICMP (Internet Control Message Protocol). Furthermore, they have the capacity to fool the attacker by simulating factors. The system response packets are suited for fingerprinting, which can be accomplished by using a tool like Nmap to scan network packets. A honeyd's attacker also uses services including Telnet, FTP, HTTP, POP3, and an SMTP (Simple Mail Transfer Protocol) server. Furthermore, they may have backdoors for viruses, such as the viruses targeted at Kuang2 and Mydoom.

B. HIGH INTERACTION HONEY POTS

In this study, we use honeynet to deploy and examine suspicious network traffic while also developing a number of tools to support our research. We provide a web interface to monitor data collection and a firewall in the backend to control outgoing connections from a possibly compromised honeypot in our architecture. Implementing a high-interaction honeypot host is a cost-effective process that is commonly employed in mid-sized organisations to gain the advantages of easy monitoring and a safe and clean successful compromise. Virtual PC, virtual box, XEN, VM ware, and user mode Linux are some of the virtual machine solutions for this environment.

We associate with a few real machines to support our production server and collaborate with low interaction honeypot zone to reach the bases and real experiment result in order to have high interaction honeypot to grant a real network information gathering and facing different scans, buffer over flows, and various analyses.

Many recent research studies have been conducted to investigate the deployment of honeypots to improve network security. Honeypots, according to Weiler's proposal, are designated as a network shield, with all incoming traffic diverted to them. Following that, information on how to disconnect that connection or how to connect legally is provided.

This technique may not be optimal because honeypots are used to lure attackers and then destroyed, rather than as a preventative or defence device. Teo presents Japonica, a solution framework with the main goal of early and rapid response to unknown threats by dynamic orchestration in detection, prevention, and reaction mechanisms to specific attacks. However, the risk of a false alarm is always a major concern until the individual personally and professionally attempts to access production systems rather than relying on Honeypots.

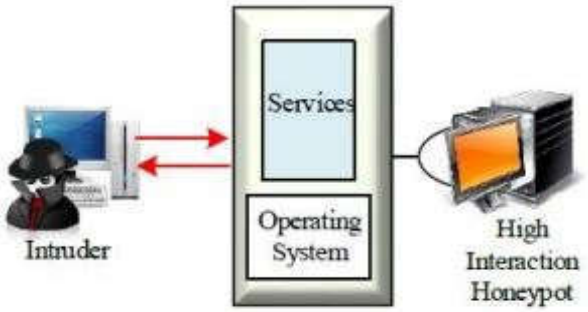

Figure 5: High Interaction Honeypots

To sum up, several of the approaches proposed above used honeypots as a defense mechanism to prevent an attacker from attacking the network. This study offered a hybrid honeypot architecture that included low-interaction and high-interaction honeypots, as well as a framework for not blocking or defending systems, but rather interactive and luring design with minimal typical faults.

C. HYBRID HONEY POTS

The demand for detailed information on a large number of IP addresses prompted researchers and network security providers to develop more intelligent and scalable systems. These studies focus on large-scale category architecture, often known as hybrid honeypot architecture. Honeypots with both high and low interactions can be used and cultivated at the same time, resulting in a cohesive unit.

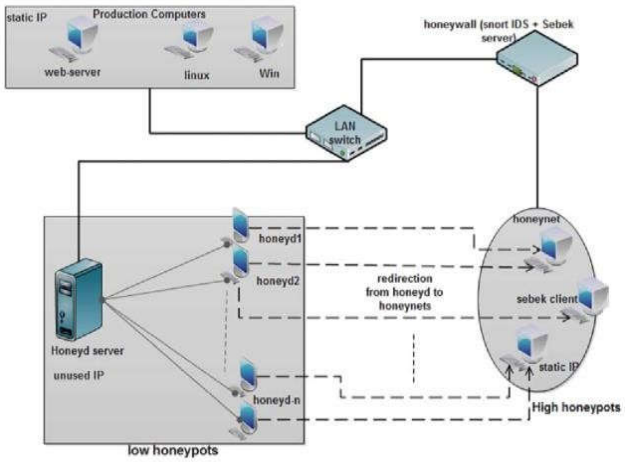

Figure 6: The proposed hybrid honeypot framework

The combination of honeypots with high and low interactions can provide the potential for deploying thousands of low-interaction honeypots onto a host, while a limited number of high-interaction honeypots are required for gathering data. Identification and obtainment of information from Internet threats can be accomplished via the different interaction levels involved in this approach.

TYPES OF MALWARE

It is helpful to classify malware in order to gain a better understanding of its tactics and reasoning. Malware is classified into numerous categories based on its goal. The following are the classes:

Virus. This is the most basic type of software. It is essentially any piece of software that is loaded and run without the permission of the user with the intent of reproducing itself or infecting (changing) other software (Horton and Seberry 1997).

Worm. This spyware looks a lot like the virus. The worm, on the other hand, can propagate across a network and replicate on other machines.

Trojan. This malware class is used to categories malware that tries to pass itself off as legitimate software. As a result, the most common spreading vector in this class is social engineering, which involves convincing users that they are downloading genuine software (Moffie, et al. 2006).

Adware. This malware's sole goal is to display advertising on the machine. Adware is frequently considered a subcategory of spyware, although it is unlikely to have drastic consequences.

Spyware. Spyware, as the name suggests, is malware that facilitates spying. Spyware's typical operations include recording search history in order to offer targeted adverts and tracking activities in order to sell them to third parties later (Chien 2005).

Rootkit. Its feature allows the attacker to gain access to data with more permissions than are permitted. It can, for example, be used to grant administrative access to an unauthorized user. Rootkits constantly mask their existence and are generally undetectable on the system, making identification and thus removal extremely difficult. Chuvakin (Chuvakin, 2003).

Backdoor. Backdoor malware is a type of malware that gives attackers a secret "entry" to the system. It has little effect on its own, but it gives attackers a larger attack surface. As a result, backdoors are never used on their own. They usually come before other sorts of malware attacks.

Keylogger. This malware class's goal is to record all of the user's keystrokes and so save all data, including passwords, credit card numbers, and other sensitive information (Lopez, et al. 2013).

Ransomware. This sort of malware encrypts all of the data on the machine and demands money in exchange for the decryption key. The user cannot open any files on a ransomware-infected machine, and the desktop picture is used to offer information about the attacker's demands.

Remote Administration Tools (RAT). This malware allows an attacker to get access to the system and make changes as if it were physically accessed. It can be described intuitively using the Team Viewer as an example, but with malicious intent.

Distributed Denial of Service (DDoS) is a type of distributed denial of (DDoS). A DDoS assault is an intentional attack that targets a website or a server in order to reduce its regular performance in a distributed computing environment. An attacker accomplishes this by utilizing many systems in a network.

5. HONEYPOT APPROCHED MODEL

A. WORMS ACTIVITY

In the context of a network, a worm is a piece of software or a programmed that, when launched on a honeypot, causes other honeypots to adjust their administration to the point where they begin to form links and create connection or pair connection requests. This distinction aids in the identification of non-self-distributing network actions from self-distributing network actions that shut the system down and configure it according to its own code. It does not, however, intend to continue the method automatically. Almost all varieties of worms contain their own executable code, indicating that the caught worms have many linkages and may have experienced a system buffer overflow or password generation. Even if most of these viable or executables have a nickname that is mostly closely associated with them, and because they are initially available as files, The accompanying Table I. shows the various worm models as well as the number of worms collected on our network.

The suggested work provides the best architecture that focuses on the best decoy the best lure architecture that can be absorbed by internal network attacks via a hybrid honey pot that can capture and record all incoming and existing data and provide us data control. The proposed honeynet records all invader activities and operations and sends them to a log for subsequent analysis.

B. DATA ANALYSING MODELE

The data analyzing module examines the information gathered from the original data. The honeynet collects data from internal honeypots and sends it to be analyzed. In the meantime, we're utilizing an appropriate firewall to gather further information about the data we've captured, and we're also sending the firewall logs to our analyzer. In the suggested architecture, a firewall module will act as a logger, capturing all traffic and its status in our back end design, allowing us to access our production systems.

C. HONEYNET ACTIVITY

The honeynet has two main functions, as previously stated: information control and information seizure or data recording. The main goal of information control is to prevent intruders from using the honeynet feature to get access to the other host. The goal of information seizure is to capture all of an invader's capabilities. It's difficult to obtain information as quietly as possible while avoiding detection by intruders. The majority of intruders attempt to encrypt channels such as SSL (Secure Sockets Layer), IPSec, SSH (Secure Shell), and other comparable channels. The encryption must be done with a specific account by the data gathering mechanism in such activities. In addition, we use seizer tools with similar capability on the honeypot to achieve a multi-record level recording method [1]. This method, not only can the many intruders' activity steps be linked together, but it can also keep the path clear of the default of a single mechanism.

The analyzing module receives logs, information, and system activity recorded by tools in the honeypot. The data is preserved as obtain data in accordance with the network connection's feature and its contents. Honeynet's recorded data is smaller in size but more dependable and deadly.

We can set up a virtual honeypot [14] on a host by utilizing the benefits of virtual technology, which is also used in honeynet. This strategy aids in the reduction and minimization of honeynet development costs. Despite this, the performance required to deploy a host is higher.

D. DE-MILITARIZED ZONE

De-Militarized Zone (DMZ) is not a network hardware device like a router or a bridge [8]; it does not pass through different packets. The De-Militarized-Zone is intended to offer secure communication with servers prior to packets entering a firewall, without the need for inbound firewall gaps between the internal LAN or network and the installed DMZ.

The policy specifies the security requirements for networks, machines, and peripherals in the DMZ. Traditional De-Militarized-Zones allow workstations behind the firewall to comment on requests destined for the DMZ. Outside the internet or public network, machines in the DMZ respond, try to forward, or reissue queries.

Many DMZs use a server (such as a proxy server) or other servers as the machines that are deployed within the DMZ. After attempting to prohibit workstations in the DMZ from initiating inbound requests, the firewall was installed. In a DMZ arrangement, the majority of the machines on an internal network or in a standard LAN run behind the firewall, allowing them to connect to an external network or the internet. To set up the secure zone, a few computers or servers are placed outside the firewall in the DMZ. These machines intercept traffic and agent inquiries for other areas of the network, and they provide an extra layer of protection for the machines inside the firewall zone.

A DMZ typically contains servers that provide various internet-based services to clients. FTP, SMTP, IMAP4 and POP3, as well as DNS server, are among the services available. Despite the fact that these servers must be connected to limited internet access, they can also safeguard the firewall. The servers and honeypots could be located in the DMZ or inside the network, although the DMZ is recommended. The best structure we are looking for that has been shown in Fig. 7.

E. PROPOSED HYBRID HONEYPOT FRAMEWORK

The proposed advancement introduces a flexible honeypot-based network security system that uses the energetic dynamic deployment and configuration of hybrid honeypots to change, in particular, organisational, financial, and vital conducted server zone networks.

The main idea behind low-interaction honeypots is to exploit free, ready-to-use IP addresses provided by operating systems or distributed systems and their services. They simulate distributed operating systems and associated services on production servers in a network. In the vast majority of circumstances, network traffic routed to honeypots will be directed to high interaction honeypots where attackers will encounter specific services. The deployment of half-breed or hybrid in order to approach honeypot technology in two main categories: Using minimal administrative hassles on account of the amount of honeyds and their specific service configurations automatically depending on the network authority Focusing on the requirement of the honeynets or high interaction honeypots in the network by the redirection of traffic scenario from the low

interaction honeypot illustrates the affinity of honyds as actual systems to attackers.

F. PROPOSED HONEYNET

The system administrator must first assign the IP addresses of the physical honeypot or essential host in the honeynet, then permit traffic redirection from low interaction honeypots, and report the activities of attackers due to the presence of phoney machines in the network. The locution redirection does not just change the route of communication between devices. It was more about reformatting the entering network packets to a certain honeyed and reintroducing them to the network. They will be able to discover their way to the true honeypot if they deploy in this manner. After that, he responds to the intruder and then provides him the idea that the invader is interacting with a real computer.

For our hybrid honeypot technique, we strive to illustrate an example of normal Local Area Network behavior. The deployment and position are depicted in Figure 7. This diagram shows a honeypot server with limited interactivity that is directly connected to the main switch and other production systems. It also depicts the actual honeypot in the architecture honeynet, which is ready to receive network direct traffic or diverted traffic from the minimal interaction honeypot. The low contact honeypots machines appear to be a physical or production system, but they are actually born as virtual machines in advance, as illustrated in the architecture.

We may employ Network Address Translation in our architecture (NAT). This method avoids the need to reconfigure each honeypot to be dynamically in internal domicile for external domiciles obtained using NTM (Network Traffic Monitoring). As a result, we should point out that by configuring the honeypot to support dynamic address reconfiguration, we can also avoid this step.

The low-interaction honeypot server depicted in this diagram has three major capabilities that are implemented in separate threads. The initial honeyd server uses a network scanning application to gather information on the various operating systems in the network, their specific direct or administrate ports, and running services, and then collects and saves this information in a file. The next thread reads the data from the file and updates the low interaction honeypot setup as needed. As a result, in the real network component, it includes the operating system, their services, and port and network support distribution. The last thread examines the honeypot log traffic data with little involvement and saves it to a certain file. Furthermore, while attackers are engaged, the servers wait for arriving traffic to flow to unused IP addresses and then suppose to identify those IPs.

A programming language, network scanning tools, and operating system must be chosen in order to build the proposed system depicted in Fig. 7. Despite the fact that the approach architecture framework is progressed in general and not limited to a certain preference. Due to the suppleness of its availability to deploy the security application, the operating system chosen for the honeyd server needed to perform open source. For this, we used the Linux Fedora 12.0 edition, which offers the required functionality thanks to our framework. The programming language required network library availability language functionality as well as the ability to easily integrate Fedora resources. In such circumstances, we use the Python programming language, which includes the required networking package.

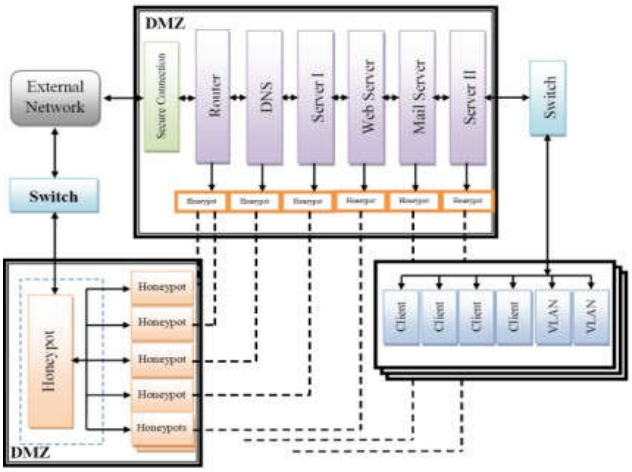

Figure 7: Hybrid honeypot architecture.

A network scanning tool was then used to determine the type of operating system used in the network as well as the various ports in the production network, and to offer the necessary information.

Nmap was picked as a component of our project for a specific purpose. This network tool can be used in two different active modes to acquire information about the many distributed operating systems that are available, as well as conduct ports and assumed running services in the network. These two folded are regular mode, which means that this tool gathered data at a particular time. In this mode, Nmap tries to parallelize port scans, even though information can be collected in the shortest time possible. However, the server may become overburdened with input or output data, and network traffic may increase as a result. The polite mode of the Nmap tool is the second mode, which gathers information slowly. The utility serializes port scans with delaying between sequential scans in this mode. This scenario applies to a machine and a network that are friendly in terms of time consumption and requiring a long time to complete scans. Regardless, we shall do a thorough network scan.

As shown in Fig. 8, the Nmap scan procedure involves issuing a ping to establish all devices on the network and collect their IP addresses, although not in a permanent file. This file can be utilized for the next scan, which will include operating system or port scans for the IP addresses that have been discovered. The results of the scans are logged into a specific file, which is often an XML file that is evaluated every time until the scan is completed. When the tool scans are completed and halted, an analyst starts and runs in a thread to extract the acquired data from the file, which creates an automatic profile to store the data.

G. DEPLOYMENT OF HONEYDS

The basic idea behind a hybrid honeypot is to use unused IP addresses, but there is a problem that helps to answer how to isolate them from the current operating system and therefore reduce the risk of revealing the real and production hosts in the network and allowing them to be hacked by intruders. A simple forward approach was implemented in order to ensure a steady continuance after integrating virtual systems into the production system via operating system distribution, and this should be done while removing physical honeypots.

6. MACHINE LEARNING APPROCHED MODEL

A honeypot is an internet platform that acts as a decoy to lure cyber attackers and detect, divert, and investigate efforts to gain unauthorised access to networks. Operating a honeypot can be costly, in part because of the specialized expertise required to design and maintain a system that aims to expose an organization's system resources while preventing attackers from gaining access to any operational systems.

Honeynet is an example of a honeypot with a lot of interaction. It can be described as a network

containing multiple systems. Honeynet can link together comprehensive information on hackers or attackers, such as keystrokes used to break into a system, chat sessions with other attackers, and the tools they employ to probe and exploit vulnerable systems. This information can reveal a great deal about the attacker.

The hybrid honeypot system combines low-level and high-level interactive honeypot systems. We employed a variety of machine learning approaches, including Random Forest, Support Vector Machine, and K-Neighbours, and finally Ensemble Learning, which is a blend of various machine learning techniques.

LOGISTIC REGRESSION

The categorization problems are solved using logistic regression. Because it can generate probabilities and classify new data using both continuous and discrete datasets, logistic regression is a key machine learning approach. Logistic regression can be used to categories observations based on many forms of data and can quickly identify the most useful factors for classification.

DECISION TREES

Assume you're working with a data set that has only one distinguishing attribute. To put it another way, one element in the data set is much more predictive of the final result than the others. We want to avoid this because averaging strongly correlated variables does not significantly reduce variance. By randomly picking attributes for each tree in a random forest, the trees become decor linked and the variance of the final model is reduced.

Every Decision Tree has some elements, such as nodes, which are the sites where it splits, and edges, which are the outcomes or results when it splits. The nodes that perform the first splits are called roots, whereas the elements that forecast the eventual outcome are called leaves.

SUPPORT VECTOR MACHINE

One of the most common machine learning issues is support vector machines (SVMs). SVMs are versatile in that they may be utilized for both classification and regression tasks.

Linear SVM is a classifier for linearly separable data, which means a dataset can be classified into two classes using a single straight line, and the classifier is called Linear SVM.

Non-linear SVM: Non-linear SVM is a classifier that is used for non-linearly separated data. This means that if a dataset cannot be classified using a straight line, it is non-linear data, and the classifier used is termed Non-linear SVM.

In n-dimensional space, there can be several lines/decision boundaries to separate the classes, but we must choose the optimum decision boundary to help classify the data points. The hyper plane of SVM refers to the best boundary. The hyper plane's dimensions are determined by the features in the dataset; for example, if there are two features (as shown in the image), the hyper plane will be a straight line. Hyper plane will be a two-dimensional plane if there are three features.

K-Neighbors Classifier

The following are the general procedures for creating a K nearest neighbor's algorithm: Calculate the Euclidean distance between the new data point x and all the existing data points in the data collection. Sort the points in the data set in ascending order of distance from x. Predict using the same classification as the majority of the K data points closest to x.

An extremely low K value, in particular, will cause your model to flawlessly predict your training data while poorly predicting your test data. Similarly, a K number that is too high will make your model needlessly complicated.

RANDOM FOREST

We implement the Random Forest Classifier using the Python computer language's Scikit Learn package and the IRIS dataset, which is a well-known dataset.The Random Forest, also known as the Random Decision Forest, is a supervised Machine Learning technique that uses decision trees to do classification, regression, and other tasks.A random subset of the training set is used by the Random Forest classifier to generate a collection of decision trees. It consists of a set of decision trees (DT) derived from a randomly selected subset of the training set, which then accumulates votes from various decision trees to get the final prediction.

ENSEMBLE LEARNING

A group of elements considered as a whole rather than individually is referred to as an ensemble. Ensemble solves a problem by creating several models and combining them.Ensemble approached said in the model's resilience and generalizability.In this post, we'll look at certain techniques and their Python implementations. Ensemble learning is the act of intentionally generating and combining many models, such as classifiers or experts, to solve a computer intelligence problem. Ensemble learning is mostly used to boost performance (classification, prediction, function approximation, etc.)

7. DOS Attack Detection using Honeypot Machine Learning

A Denial-of Service (DoS) attack is one that attempts to bring a machine or network to a halt, rendering it unreachable to its intended users. DoS attacks work by flooding the target with traffic or transmitting information that causes it to crash. The DoS attack deprives genuine users (workers, members, or account holders) of the service or resource they expected in both cases. DoS assaults frequently target high profile institutions including banks, commerce, and media companies, as well as government and trade organisations. DoS assaults rarely result in the theft or loss of sensitive data or other assets, but they can cost the victim a lot of time and money to cope with. DoS attacks can be carried out in two ways: flooding or crashing systems. Flood assaults happen when a system receives too much traffic for the server to buffer, slowing it down and eventually stopping it. A random forest is a meta estimator that employs averaging to increase predicted accuracy and control over fitting by fitting a number of decision tree classifiers on various sub-samples of the dataset. If bootstrap True (default), the sub-sample size is regulated by the max sample's argument, otherwise, the entire dataset is utilised to create each tree.

```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier()

#fit random forest classifier on the training set
rfc.fit(train_x, train_y)

# extract important feature
score = np.round(rfc.feature_importances_, 3)
importances = pd.DataFrame({'feature': train_x.columns, 'importance':score})

importances = importances.sort_values('importance', ascending=False).set_index('feature')
print("==================================Fature of DOS Attack==================================")
# plot importances
plt.rcParams['figure.figsize'] = (11,4)
importances.plot.bar()
```

Figure 8: Future of DOS Attack

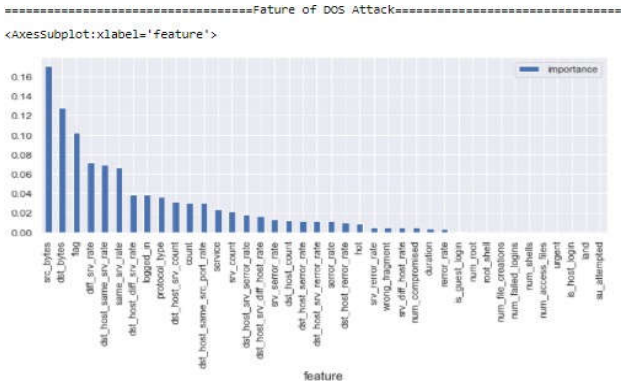

Figure 9: DOS Attack Detection

```
============================== Decision Tree Classifier DOS Model Evaluation
==============================

Cross Validation Mean Score of DOS:
 0.9960869883971739

DOS Model Accuracy:
 1.0

Confusion matrix of DOS:
 [[8245    0]
 [   0 9389]]

Classification report of DOS:
              precision    recall  f1-score   support

     anomaly       1.00      1.00      1.00      8245
      normal       1.00      1.00      1.00      9389

    accuracy                           1.00     17634
   macro avg       1.00      1.00      1.00     17634
weighted avg       1.00      1.00      1.00     17634
```

Figure 10: Decision Tree Classifier DOS Model Evaluation

We implement the Random Forest Classifier using the Python computer language's Scikit Learn package and the IRIS dataset, which is a well-known dataset.Random forest, also known as Random Decision Forest, is a supervised Machine Learning technique that uses decision trees to do classification, regression and other tasks.

A random subset of the training set is used by the Random Forest classifier to generate a collection of decision trees.It consists of a set of decision trees (DT) drawn from a randomly selected subset of the training set, which then gathers votes from various decision trees to determine the final prediction.

8. Ransomware Attack Detection using Honeypot Machine Learning.

Ransomware is a type of malware that encrypts the files of its victims. The attacker then demands a ransome from the victim in exchange for restoring access to the data. There are several ways ransomware might get access to a computer. Phishing spam-attachments that arrive in an email disguised as a file the victim should trust are one of the most popular delivery tactics. They can take over the victim's computer once they've been downloaded and opened, especially if they contain built in social engineering techniques that deceive people into granting administrative access.Other, more aggressive ransomware, such asNotPetya, takes advantage of security flaws to infect machines without the need to deceive people.There are various methods through which ransomware criminals select the firms they attack. It's sometimes a matter of timing: for example, attackers may target universities since they have smaller security teams and a diverse user base that shares a lot of files, making it easier to breach their defenses.

Honeypot is a network attached system that hackers employ as a trap to detect and investigate the methods and sorts of attacks they utilize.It operates as a prospective internet target and alerts the defenders to any unauthorized attempts to access s the information system. Large firms and organizations concerned in cyber security utilize honeypots the most. It assists cyber security researchers in learning about the many types of attacks employed by criminals and attackers. It is suspected that even the cybercriminals use these honeypots to decoy researchers and spread wrong information.

Classification Accuracy is what we usually mean, when we use the term accuracy. It is the ratio of number of correct predictions to the total number of input samples.

$$\text{Accuracy}=\frac{\textit{Number of Correction predictions}}{\textit{Total number of prediction made}}$$

The precision is the ratio tp / (tp + fp) where tp is the number of true positives and fp the number of false positives. The precision is intuitively the ability of the classifier not to label as positive a sample.

```
accuracy = cross_val_score(clf_rfeDoS, X_DoS_test2, Y_DoS_test, cv=10, scoring='
print("Accuracy of Ransomware Data Set: %0.5f (+/- %0.5f)" % (accuracy.mean(), a
precision = cross_val_score(clf_rfeDoS, X_DoS_test2, Y_DoS_test, cv=10, scoring=
print("Precision of Ransomware Data Set: %0.5f (+/- %0.5f)" % (precision.mean(),
recall = cross_val_score(clf_rfeDoS, X_DoS_test2, Y_DoS_test, cv=10, scoring='re
print("Recall of Ransomware Data Set: %0.5f (+/- %0.5f)" % (recall.mean(), recal
f = cross_val_score(clf_rfeDoS, X_DoS_test2, Y_DoS_test, cv=10, scoring='f1')
print("F-measure of Ransomware Data Set: %0.5f (+/- %0.5f)" % (f.mean(), f.std()

Accuracy of Ransomware Data Set: 0.99738 (+/- 0.00257)
Precision of Ransomware Data Set: 0.99799 (+/- 0.00324)
Recall of Ransomware Data Set: 0.99625 (+/- 0.00335)
F-measure of Ransomware Data Set: 0.99705 (+/- 0.00289)

Probe

accuracy = cross_val_score(clf_rfeProbe, X_Probe_test2, Y_Probe_test, cv=10, sco
print("Accuracy of Ransomware Data Set: %0.5f (+/- %0.5f)" % (accuracy.mean(), a
precision = cross_val_score(clf_rfeProbe, X_Probe_test2, Y_Probe_test, cv=10, sc
print("Precision of Ransomware Data Set: %0.5f (+/- %0.5f)" % (precision.mean(),
recall = cross_val_score(clf_rfeProbe, X_Probe_test2, Y_Probe_test, cv=10, scori
print("Recall of Ransomware Data Set: %0.5f (+/- %0.5f)" % (recall.mean(), recal
f = cross_val_score(clf_rfeProbe, X_Probe_test2, Y_Probe_test, cv=10, scoring='f
print("F-measure of Ransomware Data Set: %0.5f (+/- %0.5f)" % (f.mean(), f.std()

Accuracy of Ransomware Data Set: 0.99472 (+/- 0.00438)
Precision of Ransomware Data Set: 0.99270 (+/- 0.00640)
Recall of Ransomware Data Set: 0.98983 (+/- 0.00622)
F-measure of Ransomware Data Set: 0.99093 (+/- 0.00570)

R2L
```

Figure 11: Accuracy, Precision, Recall Results for Machine Learning Algorithm.

SVM

```
from sklearn.svm import SVC

clf_SVM_DoS=SVC(kernel='linear', C=1.0, random_state=0)
clf_SVM_Probe=SVC(kernel='linear', C=1.0, random_state=0)
clf_SVM_R2L=SVC(kernel='linear', C=1.0, random_state=0)
clf_SVM_U2R=SVC(kernel='linear', C=1.0, random_state=0)

clf_SVM_DoS.fit(X_DoS, Y_DoS.astype(int))
clf_SVM_Probe.fit(X_Probe, Y_Probe.astype(int))
clf_SVM_R2L.fit(X_R2L, Y_R2L.astype(int))
clf_SVM_U2R.fit(X_U2R, Y_U2R.astype(int))

SVC(kernel='linear', random_state=0)
```

Figure 12: Using SVM Classifier in Honeypot.

```
accuracy = cross_val_score(clf_voting_U2R, X_U2R_test, Y_U2R_test, cv=10, scoring
print("Accuracy of Ransomware Data Set: %0.5f (+/- %0.5f)" % (accuracy.mean(), a
precision = cross_val_score(clf_voting_U2R, X_U2R_test, Y_U2R_test, cv=10, scori
print("Precision of Ransomware Data Set: %0.5f (+/- %0.5f)" % (precision.mean(),
recall = cross_val_score(clf_voting_U2R, X_U2R_test, Y_U2R_test, cv=10, scoring=
print("Recall of Ransomware Data Set: %0.5f (+/- %0.5f)" % (recall.mean(), recal
f = cross_val_score(clf_voting_U2R, X_U2R_test, Y_U2R_test, cv=10, scoring='f1_m
print("F-measure of Ransomware Data Set: %0.5f (+/- %0.5f)" % (f.mean(), f.std()

Accuracy of Ransomware Data Set: 0.99755 (+/- 0.00228)
Precision of Ransomware Data Set: 0.93843 (+/- 0.12993)
Recall of Ransomware Data Set: 0.87335 (+/- 0.13280)
F-measure of Ransomware Data Set: 0.89369 (+/- 0.11933)
```

Figure 13: accuracy, precision, Recall, F-measure

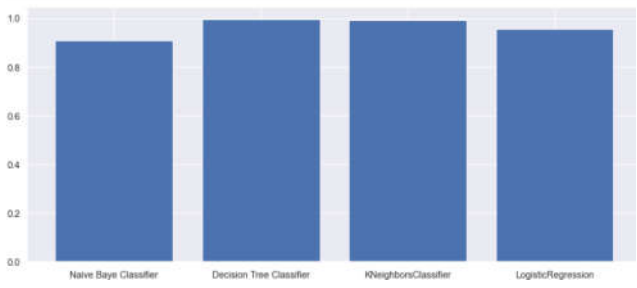

Figure 14: Comparison using Different Machine learning Algorithms

9. CONCLUSION

The suggested hybrid honeypot architecture system protects production systems to some extent. It accomplishes this by lowering the chance of hacker activity and targeting our production systems through the use of lure systems in the network, which prevent the hacker from learning about these systems, their state, or his fingerprint, and so mistaking the p honey systems for real systems. Without the redirection capabilities, we will not be able to achieve our goal, and the production system would remain exposed to direct attacks that do not travel via the controlled honeypot system. In the suggested approach, production honeypots can only perform a passive role, logging different attacker activities so that the system administrator can extract and analyze them using data mining. This could play a more active role by evaluating the attacker's activity and reducing the different types of attacks by using a signature file or a signature database that can build and my data. As we've seen, honeypots will be able to add and release warnings, as well as send notifications to the administrator, the intruder kind, and several possible solutions to stop the attack from spreading.

10. FUTURE SCOPE

1. Hybrid Honeypot System can be used for in-depth analysis of various attacks and for capturing various malware attacks.
2. Hybrid Honeypot System can be implemented in a cloud computing to reduce the security risk.
3. Hybrid honeypot system is moderate complexity and it can be used to obtain more precise information of the intruder.
4. Hybrid Honeypot System can modernize the present system and style of the new solutions for identification of user behavior.

5. Hybrid Honeypot can be made self-learning i.e. it can monitor local network and can dynamically deploys virtual honeypots based on network makeup.

11. REFERENCES

[1] Sagar Pande, Aditya Khamparia, Deepak Gupta, and Dang N. H. Thanh," DDOS Detection Using Machine Learning Technique", October 2020.

[2] Arshi M, Nasreen MD, and Karanam Madhavi," A Survey of DDOS Attacks Using Machine Learning Techniques", E3S Web of Conferences 184, 01052 (2020).

[3] Jiangtao Pei, Yunli Chen, Wei Ji, "A DDoS Attack Detection Method Based on Machine Learning", IOP Conf. Series: Journal of Physics: Conf. Series 1237 (2019) 032040.

[4] Swathi Sambangi * and Lakshmeeswari Gondi," A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression", 25 December 2020.

[5] Tuğba Aytaç, Muhammed Ali Aydın, Abdül Halim Zaim," Detection DDOS Attacks Using Machine Learning Methods", Electrica, 2020; 20(2): 159-167.

[6] DamienWarren Fernando, Nikos Komninos and Thomas Chen," A Study on the Evolution of Ransomware Detection Using Machine Learning and Deep Learning Techniques", IoT 2020, 1, 551–604; doi:10.3390/iot1020030

[7] Urooj, U.; Al-rimy, B.A.S.; Zainal, A.; Ghaleb, F.A.; Rassam, M.A. Ransomware Detection Using the Dynamic Analysis and Machine Learning: A Survey and Research Directions. Appl. Sci. 2022, 12, 172. https://doi.org/10.3390/ app12010172.

[8] Craig Beaman, Ashley Barkworth, Toluwalope David Akande, Saqib Hakak, Muhammad Khurram Khan," Ransomware: Recent advances, analysis, challenges and future research directions", 24 September 2021.

[9] Khattab M, Sangpachatanaruk C, Mosse D, MelhemR, Znati T. Roaming honeypots formitigating service-level denial-of-service attacks. In: Proceedings of the IEEE 24th international conference on distributed computing systemsMarch, p. 328–37, 2004.

[10] Krawetz N. Anti-honeypot technology. IEEE Security & Privacy Magazine, Vol. 2(1), pp. 76–9, 2004.

[11] Kuwatly I, Sraj M, Al-Masri Z, Artail H. A dynamic honeypot design for intrusion detection. In: Proceedings of IEEE/ACS international conference on pervasive services, p. 95–104, July 2004. [14] Lok Kwong Yan. "Virtual honeynetsrevisited," Information Assurance Workshop, pp 232-239, 2005.

[12] Omid Mahdi Ebadati E., Kaur H., Alam A.M., "A Secure Confidence Routing Mechanism Using Network-based Intrusion Detection Systems", OLS Journal of Wireless Information Networks & Business Information System, Open Learning Society, Nepal, pp. 83 – 93, 2010.

[13] Teo L, Sun A, AhnJ. Defeating internet attacks using risk awareness and active honeypots. In: Proceedings of the second IEEE international information assurance workshop, p.p. 155–67, April 2004. Virtual PC, http://www.microsoft.com/windows/products/win family/virtualpc/ default.mspx, 2008.

[14] Weiler N. Honeypots for distributed denial of service attacks. In: Proceedings of the 11th IEEE international workshop on enabling technologies: infrastructure for collaborative enterprises (WETICE'02) June 2002.

[15] Yeldi S., Gupta S., Ganacharya T., Doshi S., Bahirat D., Ingle R., et-al. Enhancing network intrusion detection system with honeypot. Conference on Convergent Technologies for Asia-

Pacific Region TENCON 2003, p. 1521–6, October 2003.

[16] Faizan Ullah, Qaisar Javaid, Abdu Salam, MasoodAhmad, NadeemSarwar, Dilawar Shah, and Muhammad Abrar," Modified Decision Tree Technique for Ransomware Detection at Runtime through API Calls", 1 August 2020.

[17] Samuel Egunjobi, Simon Parkinson, and Andrew Crampton," Classifying Ransomware Using Machine Learning Algorithms", Department of Computer Science, School of Computing and Engineering, University of Hudders eld, Queensgate, Hudderseld HD1 3DH, UK