e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

EXAMINATION PHASE NETWORK FORENSIC FRAMEWORK FOR IDENTIFICATION AND CORRELATION OF ATTACK ATTRIBUTES

Peter. K. Kemei*1, Dr. Moses Thiga*2, Dr. Joel Cherus*3

*1Graduate Student, Kabarak University Information Technology School Of Engineering And Technology, Kenya.

*2Senior Lecturer, Kabarak University Information Technology School Of Engineering And Technology, Kenya.

*3Forensic Security Consultant Expert Researcher, Kenya.

DOI : https://www.doi.org/10.56726/IRJMETS61000

ABSTRACT

Network forensics is a science of determining and retrieving evidential information in a computer networked environment about a criminality in such a way as to make it admissible. The established computer networks forensic field lays a strong foundation for network forensics as standard security frameworks, tools and techniques are in place for phase detection, collection, preservation and presentation of evidence. However, little has been done to address phase examination. The main challenge identified on this phase is identification and correlation. The objectives of the study were to; analyse, investigate, identify, develop and evaluate a network forensic framework which addresses the challenge in examination. A methodology was specifically formalized on real time and post attacked network traffic investigation based on datasets prototype implementation. The proposed technique in examination phase is identification and correlation of traced datasets. The identification provided attempts made in compromising a system and assist during reconstruction of intruded information. The correlation validated the particular intrusion and guide in decision to proceed with investigation. The techniques resulted in confirmation of DDoS, Portscan and cross-site scripting attacks dataset.

Keywords: Network, Forensic, Framework, Examination, Identification, Correlation.

I. INTRODUCTION

Network forensics deals with determination and retrieval of evidential information in a networked environment about a criminality in such a way as to make it acceptable. It attempts to capture traffic data logged through ingress and egress network devices. It also involves evaluation, analyses of traceable, log data of network intrusions from the current network security products, and provides information to characterize intrusion or misbehavior features. However, as we advance network in home and business, there was a need to advance network forensic view from local host to the network and web application level. There was necessity to consider these transitions into concepts, designs, models, frameworks and prototypes capabilities and implementation.

Background to the Study

Network forensics is a scientific determination and recovery of evidential information in a networked setup about an attack using an approach that makes it permissible. It has evolved to an almost established investigation process in reaction to the intruder communal and includes capturing, recording and analysing of network actions in order to determine the source of attacks according to (Sirajuddin, 2021).. The network data traffic was identified using intrusion detection systems collected from existing network security sensors tools. This data was examined for attack classification and scrutinised to trace back the intruders. The procedure can convey out insufficiencies in security products, which can be applied to guide placement and enhancement of these tools. The methodology gathers the required evidence for incident response and investigation of the crime. Network forensics expedites recording evidence for examination and assist in understanding the intruder's methodology (Mei, 2024).. It provides to understand the tools used by the intruder and new techniques in which edge defenses were evaded. (Dalal, 2021), network forensic information can also convey insufficiencies in the current sensors network security tools. These tools can be hardened to become robust

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

enough to stand the onslaught of many zero-day and hybrid attacks. There are many frameworks proposed for investigation based on network traffic attacks scenarios in network forensics. According to DFRW enhanced by (Anita, 2021) network forensics framework comprises of the seven phases as follows: "identification, preservation, collection, examination, analysis, presentation, and investigation. Examination phase deals with comprehensive organized examination of evidence linking to the alleged attack. The main challenge of examination phase is packets identified are examined to capture protocol details, which are manipulated by intruders. This information is correlated with attack events and the compromise is validated. Validation of attack takes the process to the investigation phase. Packets are reorganized into individual transport-layer connections between machines and the attack behavior are analyzed by replaying the attack.

The established computer forensic field lays a strong foundation for network forensics as standard procedures and tools are in place for detecting, collecting, preserving and presenting network evidence as opposed to examination, analysis and investigation phases still experiencing forensic challenges. The examination phase lacks effective mechanism to identify, correlate and validate the features of packets that have been manipulated by attackers.

Identification and Correlation of Network Events

The general purpose frameworks have been proposed and recommend as approaches which aim for directly supporting the inherent workflows of the network forensics domain. Most approaches operate on raw network traffic and offer search capabilities at varying granularity. Bro (2022) network security monitor offers a high-level policy for network analysis along with a rich type, event monitoring based scripting language. Bro reconstructs raw packets into transport-layer byte streams, which protocol analysers then dissect into fine-grained streams of application-specific events. User can write handlers for these events to perform arbitrary computation. Bro also ships with a library of scripts which record the protocol activity in detailed log files. In previous work, developed the Bro cluster according to Matthias (2020) to scale the analysis to multi gigabytes links. This forensic framework mode, a load-balancer dispersed the network traffic of packets over the entry and exit of network nodes devices, such that packets from the similar connection arrive at the same terminal node. Since Bro generates only network log files and does not come with a perseverance constituent, physical search rapidly runs into large scale issues by allowing a scalable output for perseverance that can natively capturing and storing specific logs files. (Gregor, 2020) the timestamp network protocols records raw network traffic and builds tree indexes for a specific packet headers fields. In order to manage stream of network traffic capacities, the network security systems and tools tracks connections allowing the packets belonging to the similar flow after the connection byte stream has attain a specific threshold. The outcome of tree-based indexes averts enough conformation of hits and evidence. For instance the querying of source and destination IP address necessitates a tree index which spans both packets fields. The design does not provide the higher dimensions of scalability. (Jihyung,2021) proposed FloSIS network forensic framework which delivers in-depth network traffic storage monitoring using the interface of the network edged devices at the granularity of stream network traffic instead of packets. The FloSIS framework capture and examine the flow stream of network traffic efficiently by implementing two stage indexing methodology. The first stage level uses two timestamp and four bloom filters to connect four tuple to connect the beginning and end to check if the flow of data contain the queried intruded information. The second stage level uses lookup logarithmic methodology to sort array of Meta base data flow. The framework exclusively exhibits form of network forensic architecture on network traffic similar to pcapIndex and time machine learning system. (Memon, 2019) proposed NetStore network forensic framework that main task implements network traffic flow archiving, stores data in flow of column format and accelerate search of specific stored data using two forms of indexing. One form of indexing implements the functionality of temporal constraints which select suitable based column using interval tree concepts. The other form of indexing implements its function using five tuple connections inverted indexes. The main challenge of NetStore framework has been identified as slow data rate of insertion of 10K record per second and revelation of query latencies of 62M records within rate of 10 seconds. This leads to low performances for figurative interactive data query at a moderate capacity data leading to non-scalable beyond the single machine system architecture deployment. Francesco (2020) proposed NET-FLi network forensic framework which utilizes single indexer network traffic NETFlow to accelerate searching of bitmap based on

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

leverages. The NETFLi framework gives optimistic result based encoding scheme which based on bit victor machine learning. The framework gives slow traffic indexing based on specific machine specification and architecture. Numerous network forensic models and frameworks have been developed to address the challenges according to Warusia (2020) and Hemdan (2021). The initiate recording of network traffic by logging information into database for inspection purposes of specific intruded attack evidence of packets. Diverse attacked features are stored which includes flow source and destination identifiers IP addresses, ports, some statistical evidence data about packets that includes packet size and packet length intervals. Altered algorithms that include network protocol analysis, apriori testing hypothesis and invulnerable that are implemented to trail network attack activities from the network logged files. These heuristic methodologies and machine learning have been adapted for modelling, examining and investigating throughout network security incidents events (Kotha, 2021) and (Qadir, 2021)). These methodologies assisting in examination phase in identification and validation correct recognition between normal and attacked packets network events (Moustafa,2019). Warusia (2020) & Soliman (2023) proposed examination converged network forensic framework that defines evidence based on digital VoIP communication . The attacked VoIP depends on variation of normal voice packets from malicious voice packets. Soliman (2023) developed a framework for digital forensics of encrypted real-time network traffic, instant messaging, and VoIP application. Ashwini (2024) proposed collecting and analyzing network model based on network evidence and network vulnerability. (Hemdan, 2021) recommended an efficient digital forensic model for cybercrimes investigation in cloud computing probablistic forensic based graph path inference evidence. The model address both false postives evidence and inspection of evidence depeding on the computation of subsequent probabilites. Aladaileh (2020) proposed a detection and examonation techniques of distributed denial of service attacks on software-defined networking controller framework implemented on forensic server. Sadegh (2020) proposed scalable platform network forensic framework that enable the forensic investigation of exploited IoT devices and their generated unsolicited activities. Also proposed framework examine and monitors suspicious patterns packets within scalable network events. Carvalho (2021) proposed network forensic based on intrusion detection systems with dynamic, static, execution and examination of malicious packets. Wang (2020) proposed hybrid network forensic framework for detecting and examining any form of network attacks mainly for machine to machine networks. The framework was design and developed for purposes of examining DDoS attacks within anti-honeypot distributed forensic system and architecture. Even though the current forensic frameworks have capacity to examine, analysis and investigate security incidents to some level, the large scale of the currents distributed systems, networks and the high speed operations has leads to great challenges when extracting essential information from these architecture. The information comprises from suspicious network events, privacy issues that discloses to what levels the risks of those security incidents. Furthermore, current network forensics frameworks takes great computational system and human resources when examining, analysing and investigating distributed networks in large scale architecture without aggregating significant network traffic streams exclusive of suspicious network security incidents.

Network Examination Phase Forensics Domain

Saeed (2023) emphasis the network technologies emerging such as block chain, big data, artificial intelligence, data analytics, the industrial Internet of Things and cloud computing are critical enablers for digital transformation. Due to extensive benefits businesses accelerating the digital transformation drive. A network security incident has grown into a significant challenge for business and to gain business continuity, organizations need to secure their digital transformation tools, methodologies and artefacts. Rachana(2022) proposed a method for applying security engineering to build security countermeasures. It identifies the threats, undesirable event characterized in terms of a threat agent, a presumed attack method, a motivation of attack, and an identification of the information or systems under attack. DDoS threats deplete the network resources rapidly particularly link parameters. Modelling these attacks provides a strong base for analysing the attack characteristics. Gupta(2021) recommended grid-aware distributed model predictive control of heterogeneous resources in a distribution network: The model is based on real-time systems, stochastic processes, processor scheduling, computational modelling, predictive control, voltage control, reactive power ADMM dispatch, distributed control, model predictive control, and power distribution network. Law

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

enforcement agencies need uses in monitoring, detecting and analysing the network traffic to when investigating network security incidents on real time bases. This may be against the goal of maintaining privacy of individuals whose network communications are being monitored. Haider(2022) proposed an enhanced interface selectivity technique that improve the quality of service for the multi-homed node. Capability based alert correlation uses notion of capability to correlate IDS alerts where capability is the abstract view of attack extracted from IDS alerts/alert. To make correlation process semantically correct and systematic, there is a strong need to identify the algebraic and set properties of capability. Fornasini(2019) proposed framework bade on observability and reconstructibility properties of probabilistic boolean networks on a finite time interval addressed state. Javed (2022) presented a comprehensive survey on computer forensics: state-of-the-art, tools, techniques, challenges, and future directions. The time interval between events promises to reveal many key associations across events, especially on multiple sources. The time interval is then used as a parameter to a correlation function that determines quantitatively the extent of correlation between events.

Examination phase challenges

Examination phase with the existing framework cannot identify useful network events for detecting the attacks. The various protocols features being manipulated by attacks need are not listed. Correlation of the attacks features with possible attacks scenarios are not performed. The attack must be identified and no validation done before decision making decision to proceed with investigation analysis. Effective mechanism are not in place for identification of attacks features from packets capture.

Identification and Correlation Framework Network Security Techniques

The framework addresses challenges associated with examination phase. The main challenges of examination phase in other framework is identification and correlation of network events with attack features. The conceptual framework addresses the challenge of identification and correlation of network events using open source applications to open the file, read the contents of the file, encode and extract various protocol features indicating each field with a self-explanatory attribute name. Identifying important sessions of suspicious activity will reduce the data to be analyzed. The correlation of events validate the occurrence of the malicious incident and guide the decision to proceed with the investigation as shown in the diagram below. The research framework for developing proposed network forensic framework for managing network security incidents is illustrated in figure 3.2**.** The examination phase input from system, networks, security system and application captures data network traffic for that has been identified and correlated by specific protocol through selection of attacked attributes and key features selection. The data traffic examined are preserved through MapReduce processing algorithm access through active storage collection systems

II. METHODOLOGY

Identification and correlation of network events main objective of developing a network forensic framework for managing network security incident particular at the examination phase. Reith(2013) recommended a practical challenging circumstance that necessitates particular practice of intrusion in imperative to progress it. This intrusion necessitates the proof of identity and examination of a specified problematic state of affairs to progress a profound thoughtful of the difficult capacity in imperative that appropriate resolution can be recognized. Important network events are identified and correlated with attack features that manipulate network protocol header fields. The header information of each packet in the capture file are read and various fields that are manipulated for attacks are examined leading to specific identification and correlation of attacked information relied from network security incidents pre-processing information. These information present more details based on representation attributes and attacked features, scaling all attributes into selected range identification of optimum features. The solution is to standardize and improving the quality of network forensic information into a unified form and normalize it into an acceptable input before presenting to subsequent examination phase for further analysis using particular examination phase forensic framework.

Research Framework

The purpose of examination phase to examine data captured and identify by multiple network protocols. In order to perform and support automated identification and correlation on the captured attacked information.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

In this research attributes and selection of attacked information are normalization and merging actions are included. The groupings of attacked information are based on the similarities of several attributes and features values using selection algorithm. Series of attack steps are revealed by identifying the amount of clusters produced. Common attack steps can be recognized by looking at the large clusters. A collective for feature evaluation, features selection algorithms were invoked to select the set of relevant features and data integration algorithm were invoked for data size reduction. The output of the proposed examination phase framework manages the attributes and attacked features, scaling all attributes into selected range, identify optimum features and incidents pre-processing that are examined resulting to identification and correlation of attacked information as shown in the figure1 below

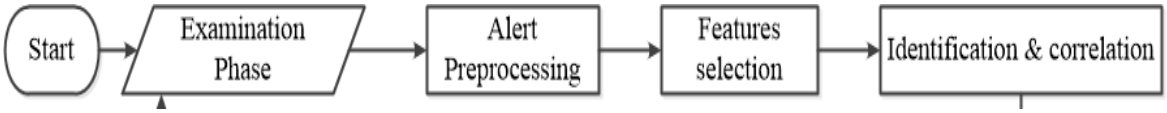

Figure 1: Conceptual examination framework flow diagram for identification and correlation of attack attributes

Saeed (2023) recommended the need to implement different methodologies and frameworks related to digital forensics investigation, incident response and elaborates the impact of forgery and tampering in the evidence chain-of-custody. The network traffic events comprises of header information of each packet in the capture file are read and various fields that are manipulated for attacks are examined leading to specific identification and correlation of attacked information relied from network security incidents pre-processing information. These information present more details based on representation attributes and attacked features, scaling all attributes into selected range identification of optimum features. The solution is to standardize and improving the quality of network forensic information into a unified form and normalize it into an acceptable input before presenting to subsequent examination phase for further analysis. The figure 2 shown below illustrates mapping of challenges of examination.

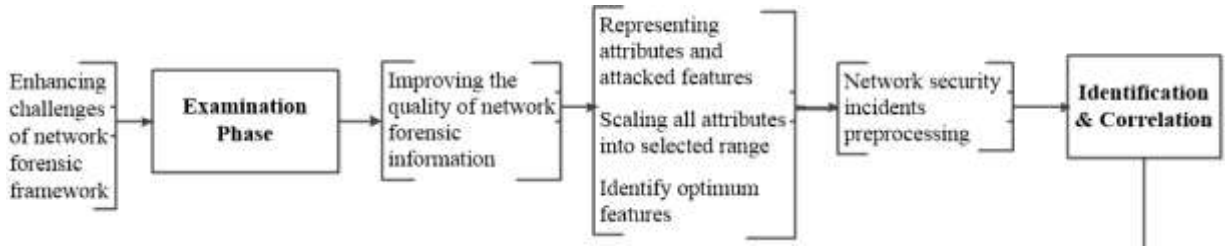

Figure 2: Flow diagram mapping challenges of examination identification and correlation of attack attributes

III. MODELING AND ANALYSIS

Examination Phase (Identification and Correlation Technique)

The main goal of the examination phase framework was identification and correlation of network events for capturing network packets from suspicious attacked network hosts and transfers the attached packets to analysis forensic database server. The details evidence captured from libcap.cap libraries contains packet header, captured number of packets, and the length of captured packet. The main libcap.cap library file used to extract packet header attributes snaplen details is perl language module Net::Pcap, which enables to capture and correlated all packet length size stored. The details information contained in the libcap.cap file for numerous packet headers such as TCP, IP, UDP and ICMP protocols were extracted from the file and stored in the forensic network database Net.Pcap server respectively. The libcap.cap enables encoding of captured packets and extracts numerous network protocols attributes that specify each packet field name. The packets encoded and extract corresponding to ver, TTL, flags, ID, offset bit, cksum, proto, srcip, destip, and options attributes. The specific associated with TCP protocol packet header attributes captured, and correlated by Net.Pcap file includes src_port, dest_port, seq_num, , ack_num, flags, hlen, flags reserved, urg, winsize and options. The specific associated with protocol packet header attached attributes captured and correlated by Net.Pcap file includes src _mac, dets_mac, ID, tos, flags, srcip, destip,TTL and options. For tracking the non-

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

stationary properties of flow identifiers, we apply the 'count' functions to determine all possible combinations of these flows as shown in figure 3.

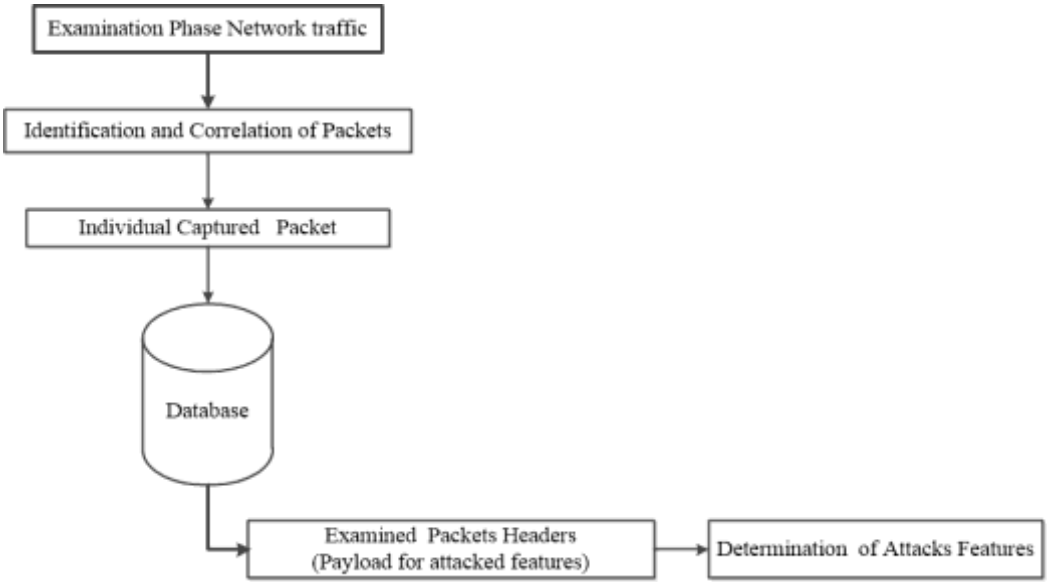

Figure 3: Examination phase framework for the identification and correlation of network events

Steps

Step (1) Read the packet captured file as inputs

The individual packet captured file have been processed by specific forensic network protocols Method are read from the database as inputs examination phase.

Step (2) Save the examined packets and experimental results. The examined and experimental results are recorded and saved in the database. It includes the details on all of the identified attacked packets and steps as well as the statistical analysis.

Step (3) The 'count' functions was applied to determine all possible combinations of these flows, as follows to examine headers (Payload for attacks features) as follows

- Select COUNT(*) as flows, srcip, dstip from network_data group by srcip, dstip;
- Select COUNT(*) as flows, srcip, srcport from network_data group by srcip, srcport;
- Select COUNT(*) as flows, dstip, dsport from network_data group by dstip, dsport, srcport;

Step (4) Save the examination and experimental phase results. The examination and experimental phase results are recorded and saved in the database. It includes the details on all of the identified and correlate of network events, examined in network and transport layer as well as in the application layer of TCP/IP model attack packets and steps used to launch network security incidents.

Description of the Network Traffic Data Sets

The UNSW-NB15 Dataset

The UNSW-NB15 dataset Moustafa at el, (2019) for research purposes in network intrusion detection systems. It is a hybrid of attack activities include real traffic and synthesized activities in a computer network traffics and comprises of nine different moderns attack types as compared to fourteen (14) attack types in NSDL-KDD datasets activities of normal traffic that were captured with the change over time (Jihyung 2021). The UNSW-NB15 dataset has forty-nine (49) features that comprised the flow based between hosts (like, client-to-server or server-to-client) and the packet header which covers in-depth characteristics of the network traffic. This data set contains 2, 540,044 observations

In UNSW-NB15 data set, there are nine categories of attacks:

1. Fuzzers: In this attack, randomly generated data is feed into a suspend program or network.
2. Reconnaissance: Attacker gathers information from the system and stimulates the attacks.
3. Shellcode: It is code used as the payload of a network packet to exploit network attacks.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

4. Analysis: This attack includes port scan, spam and HTML files penetrations.
5. Backdoors: Access of a system is gained by silently bypassing the security mechanism.
6. Denial of Service where attempts are to shut down, suspend services of a network resource remotely making it unavailable to its intended users by overloading the server with too many requests to be handled.
7. Exploits: The attacker exploits the vulnerabilities of the system through the known loopholes of the system.
8. Generic: The attack is implemented without knowing how the cryptographic primitive is implemented and works for all block ciphers.
9. Worms: The attack replicates itself to spread through the network.

Table 3.1: Attack Distribution in UNSW-NB15 Data Set

Category	Training	Set Testing set
Normal	56,000	37,000
Analysis	2,000	677
Backdoor	1,746	583
DDoS	12,264	4089
Exploits	33,393	11,132
fuzzers	18,184	6062,
Generic	40,000	18,871
Reconnaissance	10,491	3,496
Shellcode	1,133	378
Worms	130	44
Total Records	175,341	82,332

The UNSW-NB15 dataset has been divided into two Training datasets (82, 332 records) and a Testing dataset (175, 341 records) including all attack types and normal traffic records. Both the Training and Testing datasets have 45 features. The features scrip, sport, dstip, stime and ltime are missing in the Training and Testing dataset.The UNSW-NB15 data set has several advantages when compared to the NSLKDD data set. First, it contains real modern normal behaviors and contemporary synthesized attack activities. Second, the probability distribution of the training and testing sets are similar. Third, it involves a set of features from the payload and header of packets to reflect the network packets efficiently. Finally, the complexity of evaluating the UNSWNB15 on existing classification systems showed that this data set has complex patterns. This means that the data set can be used to evaluate the existing and novel classification methods in an effective and reliable manner.

Evaluation of Examination Phase

The events specific to DDoS attacks, port scan attacks and CSS attacks were used as a case study to evaluate identification and correlation phase in examination phase. The attacks cannot be classified based on a single packet information and can only be decided upon observing a sufficient number of packets. The tables are created to perform statistical analysis and calculate the thresholds for various attacks. These derived attribute values are calculated from basic attributes of a packet and help in attack detection queries

Selecting Imperative Features

Apart from selecting only substantial network traffic flow stream, the substantial features in the network traffic flow was incorporated during the network forensic framework for managing security incidents during implementation stage. We applied statistical the chi-square feature selection method x^2 (Saputra, 2022) owing to its simplicity application at real-time. X^2 was used to calculate and measure the security intrusion attacks incidents of two independent variables in relation to their class label and then the top most variables ranked were picked as significant features using equation (1).

$$x^2 = \sum_{i-1}^{y} \sum_{j-1}^{c} \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}} \quad (1)$$

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

Where x^2 refers to the chi-square of independence, O_i, represents the observed value of two variables and $E_{i,j}$ represents the mean of two variables.

The statistical thresholds were calculated from these features for numerous intrusions using statistical correntropy-variation technique equations 2, 3 and 4 which is a combination of correntropy according to (Yunfei, 2021) for measuring estimates similarities intrusion security attack instances, normal traffic and a variation threshold when identifying attacks discovers. The nonlinear correntropy is a comparison function that exposes the associations among abnormal and normal observations, whereas the variation estimating how far the abnormal instances from the normal ones. Resulting features ideals were calculated from elementary features of a packet. This evidence were used to excerpt suspicious traffic from the incarceration records. The attacked traffic was transformed back into the origin traffic internment format permitting examination via existing current open source components and tools. Identification and Correlation of protocol features, which remain influenced by the intruders, is an endless practice. Identifying the intruders from traffic internment records requirements capability in checking packet data for over a period. A datasheet of entire potential traffic and at entire layers was organised over a period. The public methods and interactive forms of the intrusions also scrutinised and correlated. This assisted during determination and examining original network intrusions security incidents.

A correntropy of two random variables (f1 and f2) is estimated using equation (2) as follows:

$$V_\sigma(f_1, f_2) = E[K(f_1, f_2)] \quad (2)$$

Where E[ ] represents the mean of the features $\mathbf{K_\sigma}$ represents the Gaussian function $\boldsymbol{\sigma}$ represents the size of the kernel computed through the following equation (3)

$$K_\sigma(.) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(.)^2}{2\sigma^2}\right) \quad (3)$$

The combined probability density function ($P_{FI,F2}(f1, f2)$) is anonymous, although a predetermined number of observations ($\{f_1, f_2\}^M$ i, j = 1) was identified. The correntropy estimate is computed using equation (3) as follows:

$$V_{M\,\sigma}(A, B) = \frac{1}{M}\sum_{i\,j-i}^{m} K(f_1 - f_j) \quad (4)$$

To carry out the experiments, we chosen random samples based on proposed redundancy algorithm from the examined dataset with some captured and saved based on timestamp integration algorithm sample based on sizes between define range limit from selected the essential features using the chi-square method and investigating attack activities using equation 1,2,3 and 4 correntropy-variation technique.

IV. RESULTS AND DISCUSSION

The intruder objective is to attack specific victim in a network setup using particular type of protocol to manipulate packet field and gain access to victim system and information compromising the confidently, privacy and integrity without the knowledge of the owner. The analysis attacked information evidence of various forms of intrusion which were identified, examined and correlate which specific packet parameters. The two types of intrusions that were analysed at the network and transport layers are port scan and Distributed Denial of Services (DDoS) respectively. The identified intrusion and correlated protocol attributes fields examined are as shown in table 1 and table 2 in subsequent section.

Table 1: Attack and Protocol Feature Correlation of Port Scan Attacks

Attack	Protocol fields to be examined
Connect Scan	Enormous amount of unsuccessful connects and successive port number requests
SYN Scan	SYN Flag and no corresponding ACK
FIN Scan	FIN Flag and Sequence number
ACK Scan	ACK Flag and RST as replies
Null Scan	No Flags set
XMas Scan	URG, PUSH and FIN Flags set

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

UDP Scan	UDP requests and ICMP port unreachable messages
Ping Sweep	Type = 8 and code = 0

Full packet capture is performed on the web server using Wireshark in application layer to evaluate cross-site scripting attacks details. The identification HTTP attributes and correlation of the payload were performed with the attack vectors. The packet capture file is opened in Wireshark and all the packets who comprise of the server's IP address as source or destination were selected. Attacker login and cookie information can be read by examining the payload of TCP ACK segments being sent from the server to the client. Attacker placing the malicious script for stealing victim's cookies were seen in the payload.

We applied the correntropy for multivariate forensic network data, as provided in equation (4), we calculate it for both normal and suspicious attacked network vectors observations as

$$I_{1N} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22}\ f_{ij} \end{bmatrix}, Y_{1N} = \begin{bmatrix} C_1 \\ C_i \end{bmatrix} \quad (5)$$

Such that I was the observations of network data, Y was the interested class label (c) of each observation, N is the number of observations and f was the number of features.

The mean of correntropy values of normal vectors (normalcorpy) was computed using equation (6) in the examination phase. In the observation phase, the correntropy value (corpytest) was estimated for each record based on equations 4 and 5.

We design a baseline between the μ(normalcorpy) and each testcorpy using the standard deviation measure(σ), which estimates the amount of variation between the mean of normal correntropy values and each correntropy of testing records. If the variation between the two values is greater than or equal (2σ), the testing vector was considered as an attack, as given in equation 7. This was because such a vector was so far from the dispersion of normal correntropy values and was difficult to fit it within the same distribution of normal data. We called this threshold a Risk Level (RL) that can identify all attack observations with low false alarm rates. The RL is scaled in a range of [0, 1] in order to exactly specify to what extent the abnormal activates deviate from normal ones.

$$\mu(\text{corpy}^{\text{normal}}) = \frac{1}{N}(\text{corpy}^{\text{normal}}) \quad (6)$$

$$RL = \begin{cases} \mu(\text{corpy}^{\text{normal}}) - (\text{corpy}^{\text{test}}) \geq 2\sigma & \text{attack normal} \\ \text{else} & \text{normal} \end{cases} \quad (7)$$

For example, Table 1 lists some flow identifiers from the UNSW-NB15 as shown in dataset with estimated RL values.

Table 2: UNSW-NB15 Features of the Proposed Examination Framework

CATEGORY	FEATURE NUMBERS
Normal	11,34,19,20,21,37,6,10,11,36,47
Dos	6,11,15 16,36,37,39,40,42,44,45
Fuzzers	6,11,14,15,16,36,37,39,40,41,42
Backdoors	6,10,11,14,15,16,37,41,42,44,45
Exploits	10,41,42,6,37,46,11,19,36,5,45
Analysis	6,10,11,12,13,14,15,16,34,35,37
Generic	6,9,10,11,12,13,15,16,17,18,20
Reconnaissance	10,14,37,41,42,43,44,9,16,17,28
Shellcode	6,9,10,12,13,14,15,16,17,18,23
Worms	41,37,9,11,10,46,23,17,14,5,13
Common	6,9,10,11,12,13,14,15,16.17,36,37,41,42,44,45

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

The computed the final results depending on the highest repeated features with at least three times. The feature vectors and flow network identifiers such as source IP (srcip), source port (sport), destination IP (dstip), destination source (dsport) and protocol types (proto) were selected using the simple random sampling technique in order to remove repeated instances or missing values, improving the overall performance of network forensic examination phase model for managing network security incidents .Features service includes sbytes, and sttl are from the Basic Feature category. Feature smean is from the Content Features category and feature ct_dst_sport_ltm is from Additional Generated Features category. The origins of attack instances can be easily tracked via correlating their flow identifiers with their estimated RL. This way will help to define the risk level of those instances. If the RL value equals one, this means that type of attacks constitutes the highest risk to an organization as it sends many flows to a specific destination such as events of DDoS attacks. But. If the RL value equals zero, this indicates this type of attacks makes the lowest risk to that organization as shown in table 3

Table 3: Selected vectors with Risk Level (RL)

srcip	Sport	dstip	dsport	Protocol	Label	RL
192.168.168.1	179	239.255.255.250	33159	TCP	0	0.23
192.168.1.6	15982	156.67.212.136	5060	TCP	0	0.11
175.45.176.3	63888	149.171.126.14	179	TCP	0	0.25
175.45.176.2	7434	149.171.126.16	80	TCP	1	0.83
175.45.176.0	15558	149.171.126.13	179	TCP	1	0.72

Discussion

Observions of abnormal records have higher RL (more than 0.5) than normal activities (less than 0.5). Ultimately, the proposed network forensic framework for managing network security incidents examination phase defines attack activities and their risk level, helping network administrators to track and report bad events that try to penetrate their network events. As established from table 4.6, from the example of select five vectors from the USNW-NB15 dataset was designed to demonstrate how the risk level was computed based on these levels were connected with their network traffic flow identifiers for examining the evidence of attack events. The simple random sampling and chi-square techniques confirm selection of the significant observations and features that replicate the patterns of appropriate and suspicious occurrences while running the proposed network examination phase forensic model for managing network security incidents as shown below.

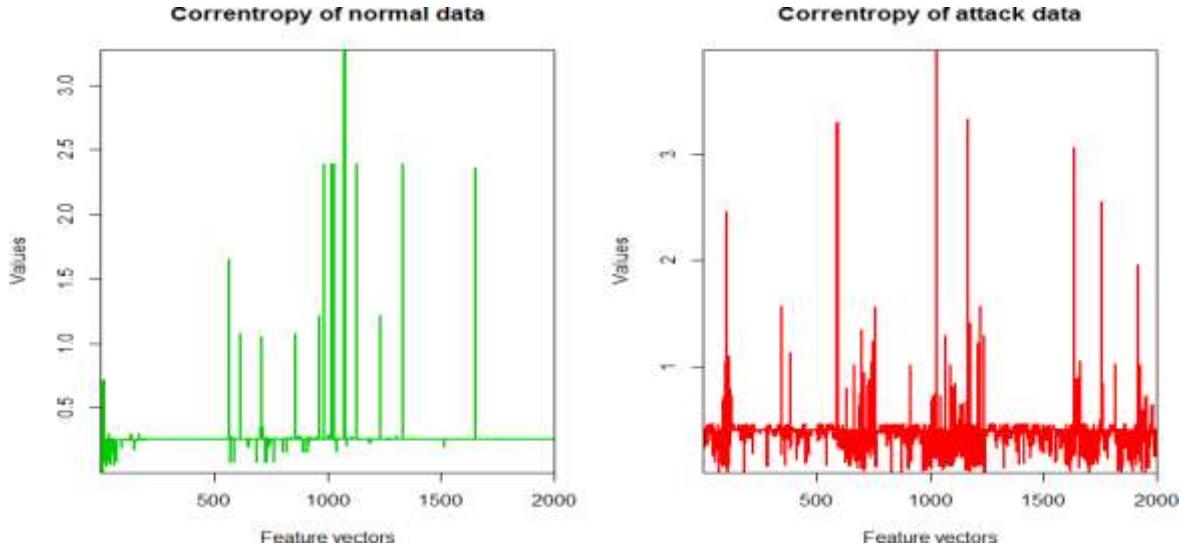

Figure 3: Correntropy of some normal and attack samples

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:08/August-2024 Impact Factor- 7.868 www.irjmets.com

The correntropy illustration graph evidently demonstrate authentic difference among normal data features and attack feature vectors as shown in figure 40 as it approximate the relationships between these nonlinear vectors. The normal samples of 2000 correntropy features flow data values clearly demonstrates difference from attack ones as illustrated in figure above. The correntropy of normal data contains even clear graphical representation as opposed to correntropy of attacks data which contain distorted randomly graphical representation. The correntropy sample between normal and attack data assist the forensic investigator to proceed in examining the attack data. As a result, the various form of network security incidents were substantially examined and explored the five flow captured identifiers which includes source IP address, source port destination IP address, destination port and type of version protocol associated with particular risk level as shown in table 3

V. CONCLUSION

In the examination phase the vital security techniques proposed were identification and correlation technique. The purpose of such security forensic techniques sequential is to improve the identification and correlation. The results show that network forensic framework produces better overall network forensic performances in comparison to other existing network forensic frameworks. This improvement is achieved through the network protocols attributes identification and correlation on attacked features

VI. REFERENCES

[1] Aladaileh, M. A.Detection Techniques of Distributed Denial of Service Attacks on Software-Defined Networking Controller. IEEE Access, Volume 3, Issue 2, August 2020,pp143985-143995.

[2] Anita, P. S. B. Roadmap of Digital Forensics Investigation. Cyber Security and Digital Forensics, Volume 5, Issue 7 June 2021, pp241-269.

[3] Ashwini, D. A. Collecting and analyzing network-based evidence. Computer Science and Information Technologies, Volume 10, Issue 2, March 2024, pp 1-6.

[4] Carvalho B. An End-to-End Framework for Machine Learning-Based Network Intrusion Detection System,. in IEEE Access,, Volume 7, Issue 8 November, 2021, pp106790-106805.

[5] Chourasiya, S. Categorizing Tracing Techniques for Network Forensics. Cyber Security and Digital Forensics, Volume 15, Issue 11, July 2024, pp 978-981.

[6] Dalal, M., Juneja, M. Steganography and Steganalysis (in digital forensics): a Cybersecurity guide. Multimedia Tools Appl, Volume 12, Issue 5, November 2021,pp5723–5771

[7] Fornasini, F. Observability and Reconstructibility of Probabilistic Boolean Networks, in IEEE Control Systems Letters, Volume 6, Issue 3, September 2019, pp 319-324,

[8] Gregor M. R. S. Enriching Network Security Analysis with Time Travel. Applications, Technologies, Architectures, and Protocols For Computer Communications , Volume 10, Issue 3, December 2020, pp17-25.

[9] Gupta, R. "Grid-Aware Distributed Model Predictive Control of Heterogeneous Resources in a Distribution Network: Theory and Experimental Validation," in IEEE Transactions on Energy Conversion, Volume 15, Issue 6, July 2021,pp1392-1402.

[10] Haider, W & Oleiw, N. S. T. An Enhanced Interface Selectivity Technique to Improve the QoS for the Multi-homed Node. Engineering and Technology Journal, Volume 1, Issue 1, August 2022, pp1006-1013.

[11] Hemdan, E. An efficient digital forensic model for cybercrimes investigation in cloud computing. Multimed Tools Appl, Volume 10, Issue 3, March 2021, pp14255–14282.

[12] Javed, A. A Comprehensive Survey on Computer Forensics: State-of-the-Art, Tools, Techniques, Challenges, and Future Directions, in IEEE Access, Volume2, Issue 9, May 2022 pp11065-11089.

[13] Kotha, S, Rani, M. S. & Subedi, B. A Comprehensive Review on Secure Data Sharing in Cloud Environment. Wireless Personal Communications , Volume 2, Issue 5, October 2022, pp2161–2188.

[14] Matthias, V. R. Scalably Stateful Network Intrusion Detection on Commodity Hardware. The NIDS Cluster, Volume 7, Issue6, July 2020, pp223-233

[15] Mei Y. A Novel Network Forensic Framework for Advanced Persistent Threat Attack Attribution Through Deep Learning,. IEEE Transactions on Intelligent Transportation Systems,, Volume2, Issue2, May 2024, pp1-10.

[16] Memon, P. G. NetStore: An Efficient Storage Infrastructure for Network Forensics and Monitoring. International Symposium on Recent Advances in Intrusion Detection (RAID), Volume 6, Issue 1, November 2019, pp277–296.

[17] Qadir, S. & Noor B. (2021). Applications of Machine Learning in Digital Forensics, International Conference on Digital Futures and Transformative Technologies (ICoDT2) Islamabad, Pakistan, Volume2, Issue1 ,June 2022, pp. 1-8

[18] Sadegh, T. E. A Scalable Platform for Enabling the Forensic Investigation of Exploited IoT Devices and Their Generated Unsolicited Activities. Forensic Science International: Digital Investigation, Volume1, Issue1 October 2020, pp1-20.

[19] Soliman A. E. S. E. (2023). A framework for digital forensics of encrypted real-time network traffic, instant messaging, and VoIP application case study. Ain Shams Engineering Journal, Volume2, Issue1, April 2023,pp 1-18.

[20] Wang R.& Ji .W."Computational Intelligence for Information Security: A Survey," in IEEE Transactions on Emerging Topics in Computational Intelligence, Volume5, Issue2 ,August, pp 616-629.

[21] Warusia Y, Cloud Forensic Challenges and Recommendations: A Review. OIC-CERT Journal of Cyber Security, Volume4, Issue3 July pp19- 29

[22] Reith, C. G. An Examination of Digital Forensic Models. International Journal of Digital Evidence, Volume 1(3) (2013), pp12-16.

[23] Saeed, S. Digital Transformation and Cybersecurity Challenges for Businesses Resilience: Issues and Recommendations . Sensors MDPI, Volume 10, Issue 10th, January (2023), pp1-23.

[24] Moustafa, N. A. . UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Military Communications and Information Systems(MilCIS), IEEE, Volume 15, Issue 9, April 2019, pp 67-74.

[25] Jihyung L. S. L. FloSIS: A Highly Scalable Network Flow Capture System for Fast Retrieval and Storage Efficiency. USENIX Annual Technical Conference, Volume 9, Issue 3, March 2021, pp 56-64

[26] Saputra, R. Z. Feature Selection using Chi Square to Improve Attack Detection Classification in IoT Network:. International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Jakarta, Indonesia,, Volume 20, Issue 17, November 2022, pp 226-232.

[27] Yunfei Z, B. C. . Broad Learning System Based on Maximum Correntropy Criterion. IEEE Transactions on Neural Networks and Learning Systems , Volume 4, Issue 6, June 2021, pp 3083 - 3097.