Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 4s (2024)

An Object Based Decision Support Protocol for Multi- Constraint Agile Digital Forensics Investigation

T M Bhraguram[1], Dr. P.S Rajakumar[2], Dr. N. Kanya[3]

[1] Research scholar, Computer Science and Engineering department, Dr. MGR Educational and Research Institute Chennai.

[2] Professor, Computer Science and Engineering department, Dr. MGR Educational and Research Institute Chennai.

[3] Professor, Head of the department IT, Dr. MGR Educational and Research Institute Chennai.

Corresponding author: T M Bhraguram (e-mail: bhruguram@gmail.com).

Article History:

Received: *18-04-2024*

Revised: *13-06-2024*

Accepted: *25-06-2024*

Abstract:

The recent scenarios expressing live digital forensics and its applications in security domain with multiple constraint-based analysis. Many digital forensic tools available in the market helps to make a good analysis of various digital forensic situations but the awareness of the situation-based decision making is still an unsolvable issue in many critical forensic cases. As the digital forensic domain increases its scope to various business and computer related industries, leading market investors and the employees are much aware about the digital crimes and its prevention. Here in this article, we proposed a new protocol which can help the decision making of sensitive digital forensic cases by the involvement of various constraint-based evaluation. A multi constraint operational system proposed here for live digital forensic and the decision support system takes an object parameter while the decision-making procedure is implemented. Thus, the protocol introduced here may carry a set of rules where the object formation and decision support constraints bind together and it may reply on the digital forensic tool integration. The characteristics of this protocol is mainly distributed for situation awareness criterion evaluation and modelled to limit the amount of data retrieval and its acquisition. categorical acquisition of digital evidences through time, duration, devices involved, wings of affect, cause of the target, hacker/attacker/affected user, kind of attack are the constraints to be solved through this prototype. The focus of this prototype is for dynamic digital forensics instead classic forensics conducted on switched off devices. This protocol supports the standard organizational formats to follow the forensic procedures in which a detailed recommendations and documentation from various organizations. The standardized formats of National Institute of Standards and Technology (NIST), International Standards Organization (ISO), Global Professional Information Community (AIIM) and American Society for Testing and Materials (now ASTM International) into considered while developing this protocol.

Categories and Subject Descriptors

[**Digital Forensic Investigation and Decision support model**]: Decision support protocol- *Object modelling.*

General terms: Digital Forensics, decision support, digital evidence

Index Terms: Investigation Analysis, Protocol

Keywords: Protocol, Object modelling, Decision support, Constraint evaluation.

1. Introduction

This Article formulated a protocol where a detailed and sustainable form of decision making during computer based forensic analysis and an object based criterion evaluation for the digital evidences to be handled better. A rule set which can always makes use of the available digital evidences and its importance with the prediction of future knowledge regarding the threats or attacks at a particular time, network flows and the theft of sensitive data regardless of its owner priority. During a sensitive and time prioritized forensic analysis, the evidences retrieved from digital devices, networks and big data in the cloud to be examined and analyzed by the help of quality software which must be equipped with proper conditions/criteria where the workload of forensic experts can be decreased. A successful analysis returns the sensitive evidence objects falls under positive modulation can be determined and sent for the decision-making purpose and thus the negative evidence object elimination done with critical time settings. This time-oriented process done through the protocol model where the evidence acquisitions and the categorical constraints sandwiched [3][5] with the rule constraints. The constraint model includes categorical acquisition of digital evidences through time, duration, devices involved, wings of affect, cause of the target, hacker/attacker/affected user, kind of attacks. The increased cost of single constraint evaluation mostly omits the less prioritized digital evidences which may then considered to be the top data content to solve the case strategy or decision making. The protocol modelled based on the evidence categorical group, considered as objects and each object model passed through the quantitative analysis thus forms the examination phase of decision making. This model capable of handling communications network-based junk data, cloud-based fog data model, also single system evidence categories.

A situation-based knowledge acquisition system adapted with his protocol always remains the time sensitive data retrieval from Internet of Things (IoT) [7][9][12]and other infrastructures like cyber physical systems [7][9][13]. Some complex network structure prevents the data retrieval process due to its vulnerability or due the manufactures restrictions. The restrictions must be cleared with proper logical analysis and domain law enforcement where a number of analysis take place in a fraction of seconds. To find the goal of the attack and its technical aspects depends on the affected systems and the sensitivity of the data, the protocol used a time awareness model where a fuzzy correlation and a machine learning approach used.

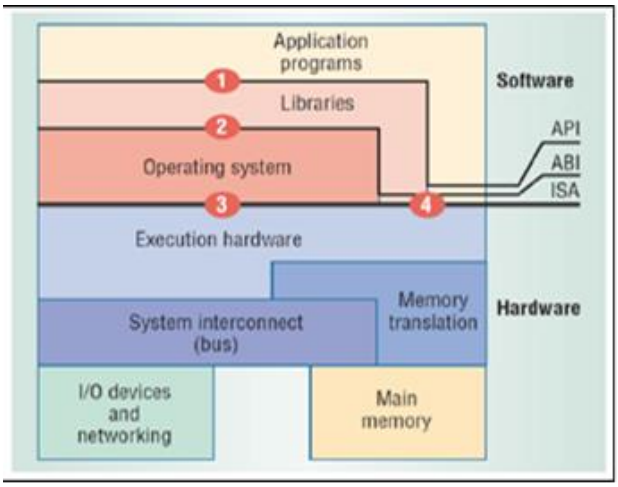

Figure 1: Digital investigation processing components

1.1 Problem Definition

The proposed model tries to make an object construction based on the categorical labels produced in a certain amount of time under restricted conditions. After finding the expert opinion and the most suitable evidence object, a further detailed examination is carried out to fulfill the evidence holes for then particular case. Most of the decision-making models failed to produce the categorical labels which are highly suitable for the specified case scenario. Situation awareness [7][4][2][8] plays a vital role to modify the category and this leads to a complicated model driven approach. Thus, the time sensitivity lost in most of the cases and the system failed to produce a time bounded result in a critical environment. This doesn't work due to the sensitivity nature of digital forensics and may change the awareness criteria also. The system can produce most critical evidence format within the time limit and carried out no lost critical constrains. Evidence pieces collected through the system analysis has no doubt on the creation or collection medium where the protocols used to extract it.

A typical classification resultant model expressed here used various classification structures and its accuracy based on then precision-recall mechanism. Each classification technique may show its accuracy level in various conditions and based on the evidence strength. Most of the file transfer mechanisms and corresponding protocols takes number of system logs to identify the precision time bound where a large volume of data transfer occurred. Multinomial neural network [2][5][10[11] reached with 95% accuracy in time critical applications and decision tree models supports 92% under the same conditions.

Classifiers	5-NN			Decision Trees (CART)			SVM (Radial Kernel)			Multinomial Neural Network			MeanDiff		
Metrics (%) / Protocols	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy
HTTP	84	80	90%	88	81	92%	84	86	91%	89	91	95%	48	100	50%
FTP	100	100		100	98		100	98		100	98		51	100	
HTTPS	78	84		82	88		83	83		91	89		0	0	
POP3	98	95		96	100		97	97		99	100		0	0	

Table 1: Protocol accuracy chart of decision-making classification

1.2 Purpose of the Model

- Develop a semantic model specification for digital forensic decision making where detailed examination is needed.
- A well-known categorical label specification for decision making [2][3][11[19] through situation awareness evidence objects.
- Identify and analyze various protocol model specification used for data transfer and obtain some relevant standards for the decision-making strategy.
- Develop a model to describe the actions or decisions made on the object constraints and incorporate any additional variables identified later in the investigation process.
- Perform a detailed examination on the most relevant evidence bunch or objects and generate a categorical label for the effective classification strategy.

- Analyze common computer or handheld device configuration properties and group into relevant associated object category.
- Analyze the crime scene situation awareness [11][12][14][16] objects where it plays critical roles under time bound.

The model exposed a categorical modulation of various time bounded situations according to the trends of the current scene. It includes a hardware device to be reexamined while the device having some operational parts during the investigation. Any persistent data storage examination can be fixed in case the device or software component has been crashed. But most of the live digital forensic situation may lead to the reexamination of specific changes on the device or software module instead a static mode of investigation strategy. After identifying a certain number of forensic indicators [5][8] affected for a threat or attack, it must be applied to the entire evidence bunch collected in a fixed time bound. The domain experts can take necessary measurements to modulate the constraints needed for the deep analysis and may correlate the specific changes on the evidence bunch by the help of protocol stack. In a live system it is most important to develop this protocol stack as it must be applicable to the various evidence bunch collected using forensic tools. The visual representation of these constraints, evidence bunch, live changes, and reexamination results are the core module of this proposal and how it can be visualized in a time bound manner.

This model gives an effective study scope of the evidence representation and its analysis through a visualized module and can be reused in various forensic tools as a supporting outcome. However, the target device during the investigation is highly sensitive on time bound and if the decision making is slow, it poses high risk due to the reason of evidence corruption or malware affection. Hence the live environment of an investigation leads to the evidence volume to be high as the tools cannot be determined in a specific period of time. The framework can be a good solution for this execution handling and may take risk assessed data to be filtered within the time frame for decision making purpose. The framework must minimize the evidence corruption score or the evidence destroying possibilities during the evidence fetching execution from the data store. During a static investigation process, the target device must be switched off while retrieving it and must reduce the possibility of malware injection [22][26] and considered as one of the most sensitive operation during the evidence collection.

During this framework execution, the system anticipates to generate a design model specification for decision making process and this design can be formulated with predefined constraint evaluation and thus the evidence and its measurements factors can be visualized using the plotting bags. It suggests a forensic tool to be used for evaluation and results of the sensitive factors in an evidence bunch with its resultant variants. These variants might be the useful decision-making hands and would no longer available in ten evidence collection records.

2. Related Works

A wide variety of forensic decision tools are available in the market with various specifications and most of them are based on static evidence retrieval without prompting any time bound specifications. Some tools are specific in dynamic retrieval with visualization module but the time constraints are limited in number as it expands according to the time. Here this article reviews short selection of recent approached related to the framework strategy.

F. Bohm, M. Vielberth, and G. Pernul. Bridging Knowledge Gaps in Security Analytics presented in 2021 [25] gives the immense security concerns while a live forensic situation deviates from its time critical evidences. The privacy beaches during the evidence selection and its preservation makes a new section criterion where deep scans and analysis plays a vital role in sensitive data modules. This might be the serious concerns in a digital forensic scene where multiple criterion to be considered as time gap, knowledge of the scene, expertness of the investigator etc. The framework proposed in this article used some critical and vital security breach concerns based on this.

F. Sommer, J. Durrwang, and R. Kriesten, “Survey and classification of automotive security attacks,” [26] made available various survey results based on the automotive attacks on evidence bunch. These automatic but sensitive and critical attacks on the evidence bunch is highly affected once the model is rescheduled for examination. This model pulled out various analysis results which may take additional categorization but take longer than usual while investigating. Additional bounds added to the existing static data but took extra time period to complete a single execution cycle may lead to the destroying of evidence bunch parts and would no longer available for the decision-making process.

M. N. Kurt, Y. Yilmaz, and X. Wang, “Real-time detection of hybrid and stealthy cyber-attacks in smart grid,” in 2019 [27] proposed a smart detection mechanism for smart grid cyber-attacks. Smart grids are always deviating to the modelled architecture format and it can take dynamic data in nature. A data exist in a smart grid can be formulated and pushed to the next cycle tunnel where a detailed excavation is expecting. But to complete a single investigation cycle and to generate an intermediate result, a smart grid is always not a good option in sensitive data evaluation. Hence the evidence bunch from a digital forensic investigation scene is critical, a smart grid may take additional parameters while considering the data.

MalViz [21] is a specialized tool developed for analyzing the behavioral characteristics of malware. Its main function is to investigate the correlation between processes within a system and an active malware. By leveraging MalViz, digital forensics specialists can promptly detect irregular patterns in process activities.

The unveiling of Event pad brought forth a novel approach to forensic malware analysis [4]. Event pad offers the advantage of streamlining network traffic samples, thereby enhancing the understanding of malware samples' networking activities. Despite its effectiveness in analyzing network data, Event pad is not suitable for detecting internal indicators for subject matter experts.

Additionally, creative methods for deciphering network packet captures (PCAPs) have been put forth [25]. Their web-based visualization design is intended to assist administrators and malware analysts in their jobs, which often entail PCAP studies. Network traffic analysis is essential for live forensics even if this technology isn't specifically made for forensic studies; nonetheless, it must be supplemented with knowledge of a device's internal communications.

Currently, the majority of visual designs used in forensic analysis focus on static analysis, a single data source, or a specific type of data like volume or network traffic statistics. Unfortunately, there is no efficient method to obtain a quick summary of a system's operations during an LDF inquiry, leaving live forensics domain experts without guidance on selecting the appropriate forensic technology. The design proposal underscores the need for a unique visual tool in the cybersecurity

field, serving as a basis for further investigation. However, there are notable limitations in the existing architecture, including being purely theoretical without evidence of practicality, particularly in terms of data accessibility. Our previous research was confined to the proposed design, rendering it non-generalizable. Defining clear requirements to direct the design and development of visual tools is not only achievable but also essential.

HOLMES [14], an innovation by Milajerdi et al., combines MITRE's ATT&CK framework with the traditional APT life cycle to create comprehensive visual representations of potential threats for analysts. To mitigate false positives, the authors have designed a TTP specification that links audit log patterns to TTPs and incorporates various noise reduction techniques.

According to references [15] and [16], they have incorporated tag propagation and policies to conduct backward and forward analysis in order to identify and construct audit log flows. However, it is important to note that the underlying assumption of the graph's formation, which assumes that every APT step can be retrieved successfully, is not supported by evidence.

The tag-propagation graph construction, as outlined by Hossain et al. [17], aims to support analysts in effectively managing the extensive amount of data. To create a concise scenario graph that accurately represents malicious activity and minimizes false positives, the authors have proposed two innovative tag propagation approaches: tag attenuation and tag decay. These approaches are based on the typical behaviors exhibited by benign Unix processes. In the field of digital forensics, graphs have been applied in a distinct manner for attribution, as demonstrated in [18] and [19]. The suggested prototype of a network forensics tool incorporates a fuzzy reasoning module that utilizes evidence graphs to establish correlations between hosts involved in the same attack.

The prototype relies on host logs as secondary evidence, with IDS alarms serving as primary evidence. In a study conducted by Studiawan et al. [20], graph-based clustering was employed to identify anomalies in access control logs. The proposed method utilizes a dynamic threshold and an enhanced MajorClust algorithm on graph structures generated from access logs to detect possible breaches, which can then be analyzed by an analyst.

Rather than directly assisting and guiding the investigator through the technical aspects of the inquiry, such as incident analysis, the mentioned methodologies aim to identify or automatically correlate specific parts of the assaults [21]. The validity of a forensic inquiry can have a profound impact on its outcome, as mishandling evidence could lead to tampering, and its applicability in supporting legal actions in a court of law. The research process must meet specific criteria, being dependable, repeatable, and well-documented. The provided resources are designed to assist investigators in effectively conducting forensic procedures in various scenarios.

Horsman [22] introduced the DERDS framework, which aids investigators in evaluating the validity of their assumptions, findings, and inferences at each stage of the investigation. Beebe et al. [23] proposed a multi-tier goal framework for digital studies that follows a similar path. Although the subsequent tiers are more complex and tailored to the investigator's decisions, the initial tier provides more general assistance. Both frameworks rely on a set of criteria given to the investigator, while our method combines cybersecurity classification with data-driven methodology. Our approach, like the one in [23], focuses on deriving future recommendations from the examination findings collected thus far.

Reasoning [24] is crucial for forensic investigators to establish logical connections between facts. A strong and clear-thinking process is essential to avoid errors in interpretations and findings. Different frameworks emphasize various aspects of investigation, with some focusing on technical analysis and others on abstraction. Case-based reasoning is a popular category that utilizes past solutions to similar issues.

SPARQLer, a forward chain rule-based reasoning system introduced by Turnbull et al. (26), is designed to assist investigators in moving from computer-specific knowledge to a more abstract level. Additionally, they presented a multi-ontology framework for digital forensics. Likewise, the authors of (27) provided a framework that enables investigators to determine whether a particular reasoning hypothesis can be applied to a system by representing the system's initial and current states as transition systems.

In their work, Nassif et al. proposed clustering strategies in [24] to tackle the task of managing a vast number of unstructured documents in forensic inquiries. Through a thorough examination of clustering algorithms, the researchers found that hierarchical methods proved to be the most optimal in terms of efficiency and outcomes.

In line with our research, De Braekt et al. [28] have introduced a model aimed at enhancing the efficiency of forensic investigations. While focuses on guiding investigators on workflow optimization and resource utilization, DISCLOSE leverages existing data and established TTPs to guide investigators through the investigation process.

3. Materials and Method

A DSS [2][4][5][11] is a model-based system for supporting and automating decision-making activities. It is composed of a model-base and a user-interface. The model-base is a set of integrated models and is the core of the system. It aids decision-makers in generating and evaluating alternative solutions to a problem (Sauter, 1997). Model-based DSS are built with model-driven development methods. This approach has several benefits to developers and may improve the acceptance of the resulting system by its users compared to classical DSS development methods. Model-based DSS development includes an explicit phase of model design, demonstrated to be advantageous for communication between developers and users, and for the overall system quality (Power, 2002). The Decisional DNA method [17[19[24] is an example of an explicit model and model-driven development process.

Decision support systems (DSS) are computer-based systems that support decision-making activities. DSS is decision-centric in nature and helps in decision-making processes. Decision support system is not a single specific system but a conceptual framework for management support systems. DSS is an interactive computer-based system that helps managers to gain insights in solving problems and take the right decision. DSS systems might serve a specific manager or a group. DSS is capable of responding to ad hoc queries and requests for data and to model it in real time. Dividing DSS [2][3][7] into components will create more understanding about DSS.

This is an obvious problem with the data file, which was evidence, due to the high rate of evidence corruption. When the case involved is an investigation regarding a company and its malpractice in computing devices, the retrieval of evidence and analysis often takes a very long time with different

phases of evidence, and some have to be done in a different period when the company is not aware. Usually, the investigation itself is done phase by phase during part-time and often changes with different cases. This type of investigation needs a systematic solution to preserve and organize evidence for easy retrieval and analysis at any time.

At present, evidence collected using certain tools can often only be analyzed using the same tool it was collected with. Should an investigation require evidence to be sent to another analyst or an analyst overseas, the tool that the second analyst can use may often be different from the tool available to the first analyst. This can lead to the evidence being converted into an alternative format to allow analysis by another tool, but often than not, covert the evidence and result in corruption. The problem becomes worse if the secondary analyst is required to testify in a case involving evidence housed in another digital forensic investigation tool.

The design of the current research is based on the selection of an object-based decision support for a multi-constraint inclusive digital forensic investigation approach to enhance the investigation process of agility in digital forensics [22][26][28]. This type of research is chosen with the aim to enhance understanding and bring valuable contributions to the entire digital forensic investigation process in both software and hardware phases. It is an effort to ensure development in designing a digital forensic investigation system with a higher criminal resolution rate, helpful to investigation agencies, and an easier approach to provide evidence in court.

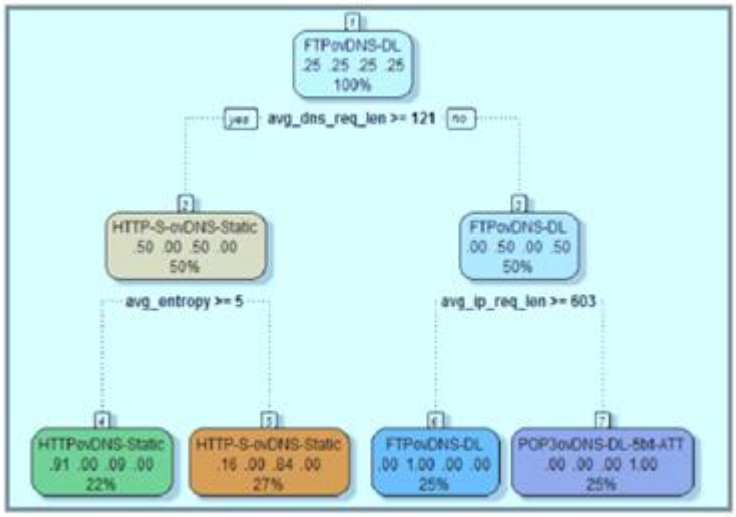

Figure 2: Protocol enforcement of IP, DNS and entropy

The methods followed as per the execution approach attached once the protocol enforcement is characterized.

Motivational search

Our primary objective is to swiftly carry out a forensic examination in order to ascertain the methodologies employed by the malicious individuals involved in the ongoing investigation. By comprehending the means through which these threat actors managed to circumvent security measures, we can gain valuable insights to prevent or minimize the occurrence of similar incidents in the future. We designate A as the comprehensive collection of potential techniques (including spear

phishing attachment, DLL search order hijacking, and drive-by intrusion) [5][6] that could have been utilized by the threat actors in this particular incident.

Occurrence Loop

To shed light on the tactics employed by threat actors in incident I, we define IY as a subset of A. For simplicity, we define IN = A \ IY as the collection of methods not utilized in incident I. The initial unknown nature of IY poses a critical challenge in cyber-forensic investigations. While certain details of IY may be known to specialists at the investigation's onset, such as the methods that initiated an intrusion detection system forwards to the breach's detection, the primary objective remains the identification of set IY. We represent the set of procedures IY as a random variable, which realization is uncertain at the beginning of the inquiry, capturing the uncertainty faced by forensic professionals. The distribution of this random variable signifies the strategies employed by threat actors, indicating the likelihood of utilizing specific subsets of tactics in a cyberattack

Past occurrences

We can establish the distribution of IY. It is important to highlight that we have access to a public repository of past occurrences, such as the MITRE Cyber Threat Intelligence Repository from MITRE Corporation. This repository includes incidents carried out by various threat actors against different targets [4][6][7][8]. Despite variations in their operational approaches, most threat actors employ similar strategies. Hence, it is plausible that certain methods frequently utilized in previous instances were also employed in the current incident. The collection of previous events is represented by the letter I. We assume that for every prior incident $\hat{I} \in I$, we possess knowledge of the set $\hat{IY} \subseteq A$, and that $\hat{IY}$ was chosen from the same distribution as IY.

State of the loop

Throughout the course of an incident investigation, forensic experts have already reviewed certain tactics, while other approaches are yet to be scrutinized. Additionally, the specialists have evaluated if the threat actors employed each method being considered during the incident. To formalize the investigation procedure, we establish $Yt \subseteq A$ as the collection of techniques examined by step t and utilized by the threat actors, and $Nt \subseteq A$ as the collection of techniques investigated by step t but not utilized by the threat actors.

State action step

It is necessary for the forensic experts to select and analyze a technique in every phase that has not been previously investigated. To formalize this decision, we define the set of potential actions in a state as the unexplored methods: in state (Yt,Nt), the set of potential actions is equal to the techniques in set $A \setminus (Yt \cup Nt)$.

Probability of evidence movement

The strategies used by threat actors are unknown to forensic professionals, causing them to treat IY as a variable with an undisclosed value at the beginning of an investigation. They evaluate whether a specific technique, a, was utilized in achieving IY. The probabilities of these transitions are then calculated.

$$\begin{aligned}&\Pr\left[\langle Y_{t+1}, N_{t+1}\rangle = \langle Y_t \cup \{a\}, N_t\rangle\right]\\&\quad = \Pr\left[a \in I_Y \mid Y_t \subseteq I_Y \wedge N_t \cap I_Y = \emptyset\right]\end{aligned} \tag{1}$$

$$\begin{aligned}&\Pr\left[\langle Y_{t+1}, N_{t+1}\rangle = \langle Y_t, N_t \cup \{a\}\rangle\right]\\&\quad = \Pr\left[a \notin I_Y \mid Y_t \subseteq I_Y \wedge N_t \cap I_Y = \emptyset\right]\\&\quad = 1 - \Pr\left[a \in I_Y \mid Y_t \subseteq I_Y \wedge N_t \cap I_Y = \emptyset\right]\end{aligned} \tag{2}$$

With the state being (Yt,Nt), Pr[a | Yt,Nt] is used to show the initial probability of the threat actors using a specific method. These conditional probabilities should be estimated based on the collection of past incidents I.

4. System Framework

The decision-support system is depicted in this article as a policy π, which links a a new action step at ∈ A \ (Yt ∪ Nt) at each time step t with a state (Yt, Nt). To computationally address the decision making module based on the objective in Equation (2), we suggest employing the k-NN and Object Based Constraint Model (OBCM) methods. In particular, k-NN is used to estimate transition probabilities, and the policy π is implemented as an OBCM algorithm.

The close monitoring of deep reinforcement learning (DRL) algorithms [13] by researchers is driven by their potential to support cybersecurity decision-making. DRL algorithms frequently encounter sample inefficiencies, demanding multiple training sessions to establish an effective policy. This poses a considerable challenge as gathering enough information from extensive datasets of past episodes is problematic. Nevertheless, rather than training a parametric machine-learning model, decisions can be made solely relying on the dataset's modest size. By directly engaging by the help of dataset, decisions would harness all the information it contains.

The Object Based Constraint Model (OBCM) tree search algorithm serves to simulate the progression of actions from a given state. It identifies the action that typically yields the highest rewards and randomly samples event sequences to determine the optimal course of action in that state. By making use of Equations (1) and (2), we can effectively utilize state-transition probabilities to replicate the cyber-forensic investigation process from any initial state.

Accurately assessing these probabilities requires an analysis of past occurrences, considering the unknown nature of the underlying probability distribution. The probability of empirical level, Pr[a | Yt,Nt], is instrumental in estimating state-transition probabilities.

$$\begin{aligned}\Pr[a|Y_t, N_t] &\equiv \Pr\left[a \in I_Y | Y_t \subseteq I_Y \wedge N_t \cap I_Y = \emptyset\right]\\&\approx \frac{\left|\left\{\hat{I} \in \mathcal{I}_{\langle Y_t, N_t\rangle} \,\middle|\, a \in \hat{I}_Y\right\}\right|}{\left|\mathcal{I}_{\langle Y_t, N_t\rangle}\right|}\end{aligned} \tag{3}$$

where

$$\mathcal{I}_{\langle Y_t, N_t\rangle} = \left\{\hat{I} \in \mathcal{I} \,\middle|\, Y_t \subseteq \hat{I}_Y \wedge N_t \cap \hat{I}_Y = \emptyset\right\} \tag{4}$$

The result is a compilation of historical events that are in line with the current status of the investigation, specifically those instances where threat actors utilized tactics from Yt but not Nt. The probability Pr[a | Yt,Nt] is calculated by equation (5) as the ratio of occurrences in I(Yt,Nt) where tactic a was used by threat actors.

This estimation is constrained by the size of the set I(Yt,Nt); as the Yt and Nt sets grow in the course of the cyber-forensic investigation, the set I(Yt,Nt) diminishes. Due to the limited number of past events, it is conceivable that none of them correspond to the current state of the investigation, resulting in the inability to calculate the estimations. To resolve this issue, we broaden the collection I(Yt,Nt) to encompass earlier incidents that closely resemble the real time investigation phase but do not precisely match it.

$$d(\langle Y_t, N_t \rangle, \hat{I}) = \left| Y_t \cap \hat{I}_N \right| + \left| N_t \cap \hat{I}_Y \right| \tag{5}$$

The count of techniques utilized in incident I but not in the preceding incident ˆI is represented by the initial term on the right-hand side, while the subsequent term signifies the count of techniques employed in the prior incident ˆI but not in incident I. Another approach involves establishing a measure known as the similarity measure, denoted as s((Yt,Nt), ˆI) = Yt ∩ ˆIY Nt ∩ ˆIN, which quantifies the instances where the current state aligns with the previous incident.

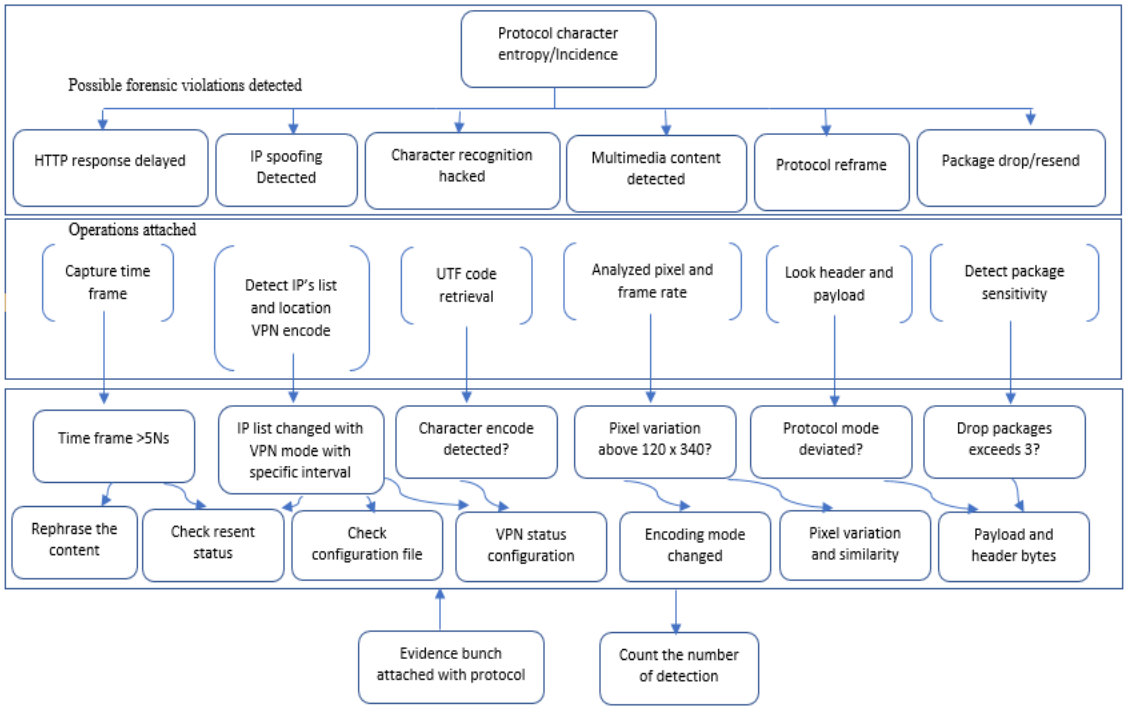

Figure 3: OBCM object flow

The system organizes the object flows in multilayer model where each layer represents the input to the next immediate layer and the intermediate results progression might be recorded with exploration decision support. The constraints like character encode and pixel variation can be exploited with fixed parameter encoding where it can vary the measurements in a fixed range. A package sensitivity is recorded with the number of dropped packages with threshold value. The payload and header values might be reconstructed after very delayed examinations. The IP detect is the coding scheme

where the location and Virtual Private Network (VPN) [13][16][19] enabled status to be encountered as it must not fall under the category of delayed examination bunch. The time frame of the captured time sequence is restricted with 5 Ns and the encoding scheme of UTF [7] is active in each session. Every session is marked under configuration file status and any changes may react with fast response time.

The exploration decision pseudocode as follows.

Algorithm 1 Exploration Decision

Function *ExplorationDecision*(Y, N, M, F):
  $MP \leftarrow \operatorname{argmax}_{\mathcal{A}' \subseteq \mathcal{A} \setminus Y \cup N,\ |\mathcal{A}'|=F} \sum_{a \in \mathcal{A}'} \Pr[a \mid Y, N] \cdot B_a / C_a$
  $n[Y, N] \leftarrow \sum_{a \in \mathcal{A} \setminus (Y \cup N)} n[Y, N, a]$
  return $\operatorname{argmax}_{a \in MP} R[Y, N, a] + M \sqrt{\frac{\ln n[Y,N]}{n[Y,N,a]+1}}$

In Equation (5), the set of k prior events closest to metric d is compared with the group of "similar" prior incidents I⟨Yt,Nt⟩, breaking ties arbitrarily. The probability measurement Pr[a | Yt,Nt] is the neighbor regression represented as K nearest, with the dataset comprising previous incidents I, the distance metric being d, and the output feature indicating whether procedure a was used in an event. While k can be fixed, our research indicates that varying k throughout the study period enhances performance. Therefore, we define k as β1 + β2 · t, with β1 and β2 as hyper-parameters determined through experimentation.

Algorithm 2 OBCM decision making Pseudocode

Input: state $\langle Y_t, N_t \rangle$, constants $\mathcal{A}, \boldsymbol{B}, \boldsymbol{C}, \gamma, K, D, M, F$
Output: action a_t
Initialization:
$\forall \langle Y, N, a \rangle : n[Y, N, a] \leftarrow 0$
$\forall \langle Y, N, a \rangle : R[Y, N, a] \leftarrow 0$
$\forall \langle Y, N \rangle : R[Y, N] \leftarrow$
  $\sum_{a \in \mathcal{A} \setminus (Y \cup N)} \sum_{j=0}^{|\mathcal{A} \setminus (Y \cup N)|-1} \gamma^j \frac{\Pr[a|Y_t, N_t] \cdot B_a / C_a}{|\mathcal{A} \setminus (Y \cup N)|}$

for *K times* **do**
  $i \leftarrow t$
  while $Y_i \cup N_i \neq \mathcal{A}$ **and** $i < t + D$ **do**
    $a_i \leftarrow ExplorationDecision(Y_i, N_i, M, F)$
    $n[Y_i, N_i, a_i] \leftarrow n[Y_i, N_i, a_i] + 1$
    if $random(0, 1) < 0.5$ **then**
      $Y_{i+1} \leftarrow Y_i \cup \{a_i\}$
      $N_{i+1} \leftarrow N_i$
    else
      $Y_{i+1} \leftarrow Y_i$
      $N_{i+1} \leftarrow N_i \cup \{a_i\}$
    end
    $i \leftarrow i + 1$
  end
  $i \leftarrow i - 1$
  while $i \geq t$ **do**
    $R[Y_i, N_i, a_i] \leftarrow$
      $\Pr[a_i|Y_i, N_i] \cdot (B_{a_i} / C_{a_i} + \gamma \cdot R[Y_i \cup \{a_i\}, N_i])$
      $+ (1 - \Pr[a_i|Y_i, N_i]) \cdot \gamma \cdot R[Y_i, N_i \cup \{a_i\}]$
    $R[Y_i, N_i] \leftarrow \max_{a \in \mathcal{A} \setminus (Y_i \cup N_i)} R[Y_i, N_i, a]$
    $i \leftarrow i - 1$
  end
end
$a_t \leftarrow \operatorname{argmax}_{a \in \mathcal{A} \setminus (Y_t \cup N_t)} R[Y_t, N_t, a]$

Throughout the investigation process, we utilize OBCM Algorithm 2. This algorithm selects an action at that is projected to yield the highest possible occurrences. By running multiple iterations, the algorithm estimates the expected number of occurrences for each action by simulating a series of states and actions from the current state, representing a random sample of potential investigation outcomes based on specific actions. In the reverse propagation phase of each iteration, the algorithm enhances its predictions of expected rewards using the information gathered from the sampled sequence. Our algorithm established the mechanism of shared exploration decision concepts.

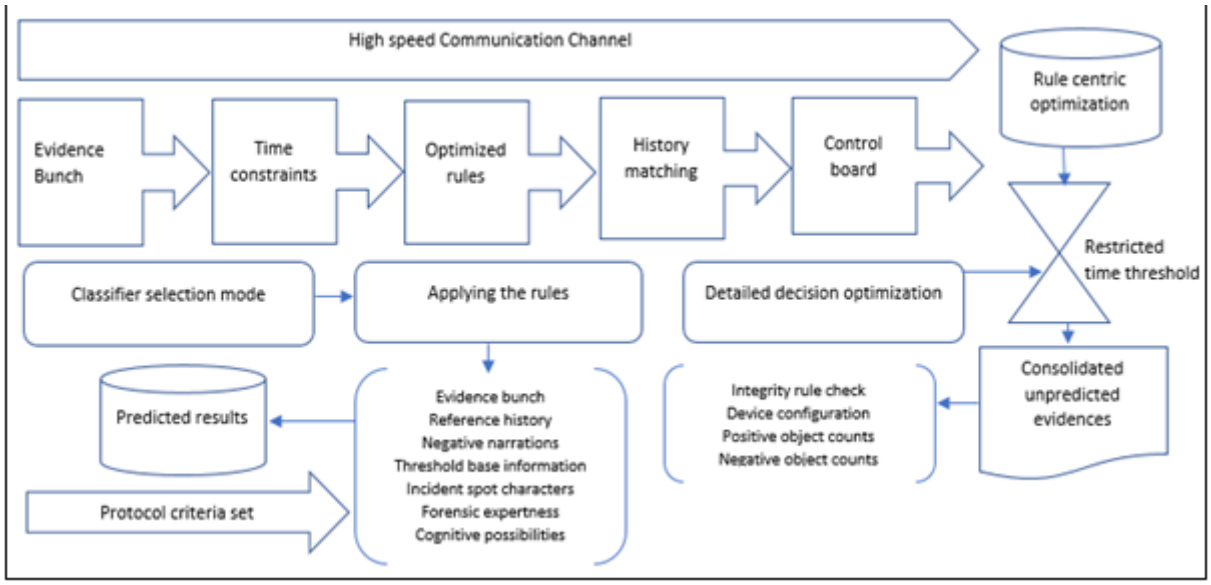

Figure 4: OBCM decision making framework

The main objective of the proposed OBCM model is to streamline the process of gathering digital evidence, reduce the amount of data extracted from an electronic machines or device parts for examination, and prioritize the investigation level of the digital forensic investigation process. The examination and analysis procedure begin with a copy of the confiscated device, which serves as the starting point. This model utilizes appropriate forensic tools and established protocols to uncover digital evidence. In this context, we demonstrate the utilization of our model during the inspection and analysis phase, which represents the fourth stage of a forensic inquiry.

The forensic specialist was aware of the cogenerated energy information system (CEIS) [22][23][25], which was implemented on the confiscated HDD [25][27][29]. The primary purpose of the installed CEIS is to monitor and control the quantity of cogenerated energy produced. However, for some reason, the system failed to provide accurate data or function properly, potentially indicating fraudulent activity aimed at concealing the actual amount of cogenerated energy produced. This theory was proposed by the specialist. Consequently, a deliberate act of suspicion was carried out against the CEIS, resulting in the manipulation of log data and potentially OS traces as well.

Operating on a high-speed communication channel and an active database with a multi-constraint rule-centric optimization channel, the OBCM decision-making framework ensures efficient performance. Throughout the evaluation process, a series of evidence bundles with time constraints are transmitted, potentially resulting in optimized rules. The historical evidence bundles are compared with the evaluation intermediate level result, and a control board actively monitors this

activity. Each evaluation is subject to a time-restricted threshold value, enabling the corresponding classifier mode to apply a range of rules. The criterion group encompasses metadata from various evidence bundles, reference history, negative narratives, threshold-based information, incident spot characters, forensic expertise, and cognitive possibilities. The framework then applies these rules to device configuration and determines the counts of positive/negative objects. Unforeseen evidence is consolidated and progresses towards detailed decision optimization, guided by a threshold value.

5. Results and Discussions

After conducting a numerical evaluation of this proposed method using publicly available datasets of actual cyber incidents. We model how our method would have prioritized the investigation for each cyber incident and depict the value of discovery as a function of effort expended.

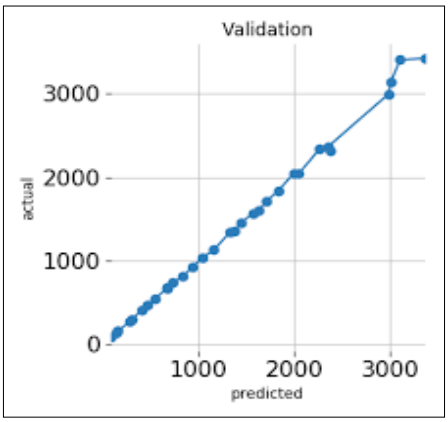

Figure 5: Cumulative benefit obtained as a function of cumulative Effort cost

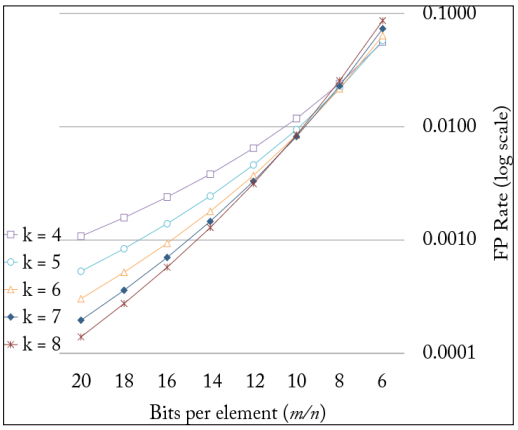

Figure 6: Evidence bunch processed in a time frame

Duration of Operation:

The OBCM method produces an average 5 seconds response time to complete for a single decision when running on a 2.4GHz Intel Core i9 CPU with a single core, and less than a second when running on many cores. These running times are small in comparison to the number of minutes that forensic analysts require to investigate the chosen technique.

A cumulative benefit validation results obtained here with actual and predicted evidence bunch and according to the accuracy measurement it provides a drastic accuracy improvement while increasing the precited and actual bunch number. Figure 6 provides the evidence element processed in a time frame and it deviates from log scale values ranges from 0.0001 up to 1.000. as the bits per element show cases the values range from 20 to 6 and it leads the coefficient values to be varied from 4 to 8. The results show the bits per element processed is inversely proportional to the coefficient values as it directs evidence processing speed up to be maximum.

While check the probability measurements a content ranking has been formulated in the results of figure 7 and it sates ten probability ranges from 0 to 0.9. Content ranking can be varied based on the evidence bunch to be processed and it gives highest probabilistic options while increasing the content ranking as its directly proportional.

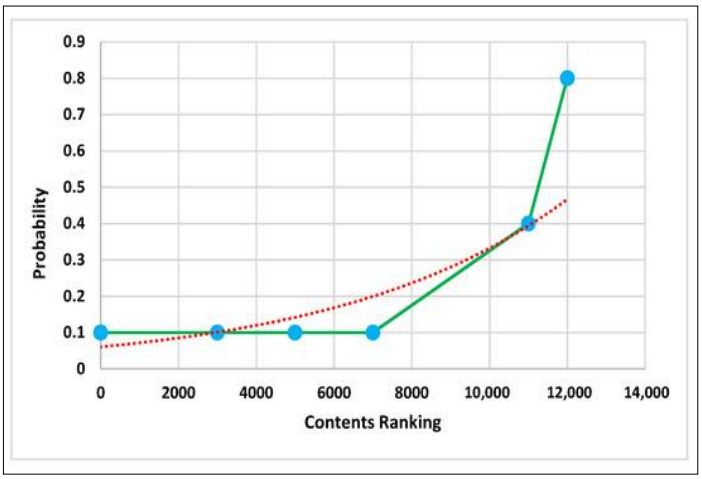

Figure 7: Highest probabilistic evidence bunches ranking

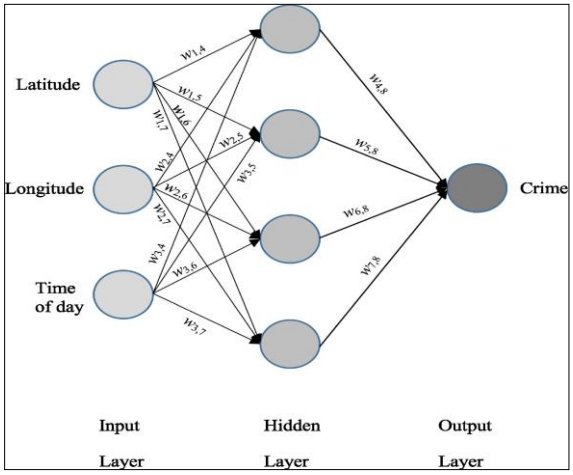

Figure 8: Node level decision making

Then decision making process engages the layer mode options in the framework and it designates various longitudinal and latitudinal measurement considerations. This could be the highest number of evidence bunch alterations and the number of input layer as depends on the time frame. A hidden layer with multiple constraint node produces the intermediate executions with optimized decision level nodes. The crime is sealed with the hidden layer node maximizations and it may lead to the possibility channels. As shown in the figure 8, the decision making process can be considered as an iterative one and it shows a single evidence bunch formulation with various measurement constraints.

6. Conclusion

To guarantee the proper application of the analysis techniques and, thus, to provide greater weight to the conclusions that may be drawn from the analysis, a tool to support the decision-making process during the forensic analysis must be developed. This is particularly important when a case calls for specialized knowledge that needs to be presented before a court. The forensic analyst and the methods employed are always being updated due to the rapid advancements in technology. In this way, it becomes clear that using flexible tools and standardized processes to support the decision-making process is essential to ensuring that this ongoing updating process is carried out appropriately.

A forensic analysis team's resources are frequently restricted, making it difficult for them to look into accidents and quickly and efficiently reduce the damage to the organization's assets. To handle the complexity resulting from the introduction of new hostile strategies and technologies, analysts must continuously improve their abilities and broaden their knowledge and methodology. To assist the analyst in overcoming these obstacles, a novel decision support model was proposed in this article. Our model is built on the development of an attack action space using forensic techniques and resources, as well as publicly available knowledge bases of Mitre ATT&CK TTPs.

We proposed an OBCM k-NN regression based computational technique, and modeled cyber-forensic inquiry as object based decision processes of cyber incidents, in order to solve the

shortcomings of methods like DISCLOSE. Our suggested approach has the important benefit of working directly with the data, rather than building a parametric machine-learning model. This means that probabilities are computed at each step and iteration based on the dataset of all previous instances. Our goal is to extract as much "mileage" as possible from the data by using non-parametric machine-learning models, which is both made possible and required by the paucity of publicly available information regarding cyber attacks. The tree search's ability to approximate best estimates and, consequently, optimal decisions based on the dataset (as the number of iterations increases) is another important benefit of our suggested methodology. Although these benefits allowed our method to surpass benchmarks, such as DISCLOSE, we discovered that the prioritization issue was extremely difficult.

REFERENCES

[1] Arshad, H.; Omlara, E.; Abiodun, I. O.; and Aminu, A. 2020. A semi-automated forensic investigation model for online social networks. Computers & Security, 97: 101946.

[2] Atefi, S.; Panda, S.; Panaousis, M.; and Laszka, A. 2022. Principled data-driven decision support for cyber-forensic investigations. arXiv preprint arXiv:2211.13345.

[3] Barr`ere, M.; Steiner, R. V.; Mohsen, R.; and Lupu, E. C.2017. Tracking the bad guys: An efficient forensic methodology to trace multi-step attacks using core attack graphs. In 13th International Conference on Network and Service Management (CNSM), 1–7. IEEE.

[4] Biasini, N. 2019. It’s alive: Threat actors cobble together open-source pieces into monstrous Frankenstein campaign. Cisco Talos Intelligence Group, https://blog. talosintelligence.com/frankenstein-campaign/. Accessed: 2022-08-15.

[5] Hossain, M. N.; Sheikhi, S.; and Sekar, R. 2020 Combating dependence explosion in forensic analysis using alternative tag propagation semantics. In 41st IEEE Symposium on Security and Privacy (S&P 2020), 1139–1155. IEEE.

[6] Hossain, M. N.;Wang, J.; Sekar, R.; and Stoller, S. D. 2018. Dependence-preserving data compaction for scalable forensic analysis. In 27th USENIX Security Symposium (USENIX Security 2018), 1723–1740.

[7] Kurt, M. N.; Ogundijo, O.; Li, C.; and Wang, X. 2018. Online cyber-attack detection in smart grid: A reinforcement learning approach. IEEE Transactions on Smart Grid, 10(5): 5174–5185.

[8] MITRE Corporation. 2022. MITRE Cyber Threat Intelligence Repository. https://github.com/mitre/cti, accessed on August 25th, 2022.

[9] Nisioti, A.; Loukas, G.; Laszka, A.; and Panaousis, E. 2021a. Data-driven decision support for optimizing cyber forensic investigations. IEEE Transactions on Information Forensics & Security, 16: 2397–2412.

[10] Nisioti, A.; Loukas, G.; Rass, S.; and Panaousis, E. 2021b. Game-theoretic decision support for cyber forensic investigations. Sensors, 21(16): 5300.

[11] Saeed, S. H.; Arash, H. L.; and Ghorbani, A. A. 2020. A survey and research challenges of anti-forensics: Evaluation of game-theoretic models in simulation of forensic agents’ behaviour. Forensic Science International: Digital Investigation, 35: 301024.

[12] Satvat, K.; Gjomemo, R.; and Venkatakrishnan, V. 2021. Extractor: Extracting attack behavior from threat reports. In 6th IEEE European Symposium on Security and Privacy (EuroS& P), 598–615. IEEE.

[13] Tong, L.; Laszka, A.; Yan, C.; Zhang, N.; and Vorobeychik, Y. 2020. Finding Needles in a Moving Haystack: Prioritizing Alerts with Adversarial Reinforcement Learning. In 34th AAAI Conference on Artificial Intelligence (AAAI), 946–953.

[14] Wei, Y.; Chow, K.-P.; and Yiu, S.-M. 2021. Insider threat prediction based on unsupervised anomaly detection scheme for proactive forensic investigation. Forensic Science International: Digital Investigation, 38: 301126.

[15] Neumann C, Champod C, Puch-Solis R, Egli N, Anthonioz A, Bromage-Griffiths A, (2007), Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae, J Forensic Sci, 54-64.

[16] Weyermann C, Marquis R, Delaporte C, Esseiva P, Dujourdy L, Lock E, Aalberg L, Dieckmann S, Zrcek F, Bosenko J (2008), Drug intelligence based on MDMA tablets data: (1) Organic impurities profiling. Forensic Science International 177 (1):11-16.

[17] Marquis R, Weyermann C, Delaporte C, Esseiva P, Dujourdy L, Koper C, Aalberg L, Dahlenburg R, Zrcek F, Bosenko J (2008) Drug intelligence based on MDMA tablets data: (2) Phusical characteristics profiling. Forensic Science International 178 (1): 24-39.

[18] Bayes T, Price, R (1763). An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S. Philosophical Transactions of the Royal Society of London 53: 370–418.

[19] Cole, S. Forensics Without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification. Law, Probability and Risk 2009, 8 (3), 233–255

[20] Koehler, J. J., & Macchi, L. (2004). Thinking about low-probability events. An Exemplar-Cuing theory. Psychological Science, 15, 540–546. http:// dx.doi.org/10.1111/j.0956-7976.2004. 00716.x

[21] Evett, I. W.; Lambert, J. A.; Buckleton, J. S. A Bayesian Approach to Interpreting Footwear Marks in Forensic Casework. Sci. and Justice 1998, 38 (4), 241–247.

[22] Skopik, Settanni, & Fiedler (2016). A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing. Computers & Security, 60, 154–176. doi.org/10.1016/j.cose.2016.04.003.

[23] Thompson, W. C. (2012). Discussion paper: Hard cases make bad law: Reactions to R v. T. Law Probability and Risk, 11, 347–359. http://dx .doi.org/10.1093/lpr/mgs020

[24] Cankaya EC, Kupka B. A Survey of Digital Forensics Tools for Database Extraction. In Future Technologies Conference; 2016; San Fransisco: IEEE. p. 1014-1019.

[25] F. B¨ohm, L. Englbrecht, and G. Pernul. Designing a Decision-Support Visualization for Live Digital Forensic Investigations. In A. Singhal and J. Vaidya, eds., Data and Applications Security and Privacy XXXIV, vol. 12122, pp. 223–240. Springer International Publishing, Cham, 2020. Series Title: Lecture Notes in Computer Science. doi: 10.1007/ 978-3-030-49669-2 13

[26] M. Beran, F. Hrdina, D. Kouril, R. Oslejsek, and K. Zakopcanova. Exploratory Analysis of File System Metadata for Rapid Investigation of Security Incidents. In 2020 IEEE Symposium on Visualization for Cyber Security (VizSec), pp. 11–20. IEEE, Salt Lake City, UT, USA, 2020. doi: 10.1109/VizSec51108.2020.00008

[27] V. T. Nguyen, A. S. Namin, and T. Dang. MalViz: an interactive visualization tool for tracing malware. In Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis, pp. 376–379. ACM, Amsterdam Netherlands, 2018. doi: 10.1145/3213846.3229501

[28] T. Wu, F. Breitinger, and S. O'Shaughnessy. Digital forensic tools: Recent advances and enhancing the status quo. Forensic Science International: Digital Investigation, 34:300999, 2020. doi: 10.1016/j.fsidi. 2020.300999

[29] A. Ulmer, D. Sessler, and J. Kohlhammer. NetCapVis: Web-based Progressive Visual Analytics for Network Packet Captures. In 2019

[30] IEEE Symposium on Visualization for Cyber Security (VizSec), pp. 1– 10. IEEE, Vancouver, BC, Canada, 2019. doi: 10.1109/VizSec48167.2019. 9161633

Bhragram T M received M.Sc degree in Computer Science and Engineering (CSE) from Bharathidasan University, India and M.E degree in Computer Science and Engineering (CSE) from Vinayaka missions University, India. His distinguished career spanning 15 years of academic and one year of corporate experience. He has published more than 19 plus articles which include Scopus & Web of Science (WoS). He holds multifarious memberships from the prominent professional bodies specifically from IEEE, ACM, ISTE, IACSIT, CSTA, IAENGG, IAHFP and IARCP. His research interest and areas are Data mining, Cloud based services, Digital forensics, Metaverse and Block chaining. He has won many awards and accolades during his career and presently pursuing as a Research scholar in the Department of CSE at Dr M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India.

P. S. Rajakumar received M. Tech degree in Computer Science and Engineering from Dr M.G.R. Educational and Research Institute (University with Graded Autonomy Status), Chennai, India and a PhD degree in Computer Science and Engineering from the renowned Jawaharlal Nehru Technological University, Hyderabad, Telangana, India. He commenced his academic career as a Lecturer and elevated to Assistant Professor, Associate Professor and Professor with a distinguished career spanning 24 years of academic and two years of corporate experience. He has published more than 20+ articles which include Scopus & Web of Science (WoS). He has also published 3 books with ISBN. He holds multifarious memberships from the prominent professional bodies specifically from CSI, ISTE, IEEE, MIEI and MTSI. His research interest and areas are Data mining, SVM classifier, IoT, cloud computing, DL and ML. He has won many awards and accolades during his career, and apart from his academic stint he has won the best NSS officer award for his yeoman services rendered and honored by the former Governor, His Excellency Surjit Singh Barnala, Tamil Nadu, India, for his relentless service and awareness towards the downtrodden society at large. Dr P. S. Rajakumar is presently working as a Professor in the Department of Computer Science and Engineering at Dr M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India.

N. Kanya received M.Sc degree in Information Technology from Alagappa University, India and M.Tech degree in Computer Science and Engineering (CSE) from Dr. M G R University, India. Her distinguished career spanning 16 years of academic and 3 years of corporate experience. She has published more than 14+ articles which include Scopus & Web of Science (WoS). She holds multifarious memberships from the prominent professional bodies specifically from IET, and IAENG. She holds 2 Indian and one international patent. Her research interest and areas are Big data analysis, IoT security, Machine learning and modelling techniques. She has won many awards and accolades during her career and presently working as Addl Dean and Prof. and head of Information technology department at Dr M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India.