INFORMATION DRIVEN SUPPORT FOR OPTIMIZING CYBER FORENSIC INVESTIGATION IMPROVED WITH SECURITY

Aparna L, Madhumitha S , Preethi K, Mrs. Sangeetha Krishnan

[1]aparnanarayanan1601@gmail.com *Student, Department of CSE, Panimalar Engineering College*

[2]madhu012012012@gmail.com Student, Department of CSE, Panimalar Engineering College

[3]kpreeth28@gmail.com Student, Department of CSE, Panimalar Engineering College

[4]sankrish2007@gmail.com Assistant Professor, Department of CSE, Panimalar Engineering College

Abstract—A web-based tool is part of the project dubbed "Information-Driven Support for Optimizing Cyber[2] Forensic Investigation improved with security." This software has the ability to confirm criminal offenses, issues, and losing people to DIG. This software offers the ability to record online crimes, make online complaints, and display missing persons information[15] on a criminal list. Any member of the public may file a complaint online. Each user logs in to the server initially to share their availability. Creating a mission statement that includes the primary duties of the unit, whether they be forensic analysis, evidence gathering, or high-tech crime investigations, is an efficient method to start this process. Cyber forensic, however, includes methods to look into or gather the data. It is described as the procedures and equipment utilized in inquiries and the collection of proof. Certain instructions, such as category-based ones, will be offered by default. Processes will be optimized to shorten the investigation process by analyzing the investigation report.

Keywords—*Cyber forensic Investigation, Encryption and Decryption, crime detection, digital forensics, Web application*

I. INTRODUCTION

The field of solving crimes has been completely transformed by modern science and technology, which has also sped up and improved the procedure. The term "forensic" refers to all of the technology and science utilized to solve crimes. This forensic management system's goal is to organize the massive amounts of data generated by the use of cutting-edge technology and scientific methodologies to solve crimes. The system will be able to save particular information in categories when generating a new case file The web application " Information-Driven Support for Optimizing Cyber Forensic Investigation improved with security " was developed with security in mind. This software offers the ability to confirm criminal acts and offers the ability to report online crimes, online complaints, missing persons show criminal[18] list and details on the web page, as well as to report online crimes. Online forums allow for public complaints.

II. RELATED WORK

There are now two categories of frameworks for supporting digital forensics: (i) attribution & correlation, (ii) forensic rating stability. In the sections that follow, we discuss the most significant earlier point out how DISCLOSE differs from or develops upon each piece of work in each area. Just the works that are most relevant to our methodology are included in this section. The reader can continue reading by referring to [10] through [12] and [13].

A. Attribution & Correlation

In order to recreate malevolent actions, a technique[20] known as event correlation involves collecting and heterogeneous data analysis from a number of sources. Event correlation is different from traditional intrusion detection in that it doesn't just look for indicators of probable harmful activity. Milajerdi et alHOLMES .'s is a system that creates elevated graphs of potentially dangerous occurrences for the investigator utilising the MITRE ATT&CK and the conventional APT life-cycle[1] architecture. The authors have created a TTP standard to map audit log flows to TTPs in order to lessen the number of false positives. Label replication and backward and forward analysis policies have been combined[10] in line with the generation of audit log flows and their detection[9] and combination. The major motivation behind the graph's designIt is unreasonable to assume that all of the APT phases will be successfully recovered. Moreover, it loses the opportunity to utilise any extra knowledge in the By focusing just on the TTPs in ATT&CK, MITRE ATT&CK[1] knowledge base. Contrarily, DISCLOSE makes use of the TTPs offered by ATT&CK[1], but it also makes use of the MITRE ATT&CK STIX[1] repository to extract pertinent information and find the probabilistic connections between those TTPs, as will be discussed in more detail later. In order to help the analyst manage the massive amount of data, Hossain et al. [17] have suggested a tag diffusion chart creation. To create a brief scenario graph of harmful conduct and lessen the quantity of false positives,

Focusing on normal Unix process behaviour, the authors have created two novel tag propagation mechanisms, decay and tag reduction. A new strategy to utilizing graphs for attribution in digital forensics was presented in [18] and [19]. The recommended mockup of a digital forensics tool utilizes a fuzzy reasoning module based on proof networks to link hosts involved in the same attack. The prototype's main display is through IDS alarms, with host logs acting as a backup. Studiawan et al. [20] employed chart grouping to find irregularities in identity management records.The recommended approach uses a dynamic threshold to discover likely violations and an improved MajorClust algorithm to access log-generated graph structures. An analyst can then review these results. Instead of assisting and guiding the investigator through the technical component of the inquiry, i.e., the examination of the occurrence, the above approaches seek to detect or automatically connect parts of the attacks.

B. Forensic Soundness

A technical inquiry's legality can have a consequence on both the inquiry's outcome—mishandling evidence might cause it to be damaged—and the ability to utilise the investigation's findings, i.e., for judicial charges in a legal proceeding. The research process must adhere to a set of rules and be trustworthy, repeatable, and well-documented [21]. The works listed below are intended to help the investigator use forensic sound approaches in a variety of cases. Protocol for Digital[8] Proof Disclosure and Assessment, put forth by Horsman [22], aids the investigator in evaluating the validity of her deductions, assumptions, and conclusions at each stage of the inquiry. Beebe et al. [23] developed a similar strategy as an outline for number of co technological study. While the rest are more intricate and tailored to the preferences of the investigator, the first tier offers more general assistance. While our technique uses a both systems are constructed according to a variety of criteria that are given to the investigators, and they combine data-driven technique with vulnerability categorization. Our strategy and that in [23] have one thing in common: we are basing our recommendations for the future on the findings of the current inspection[16]. A more technical approach has been proposed in [24] to guarantee the authenticity of an inquiry. The authors' conclusion has been reached. a system that can detect anti-forensic assaults and correct their effects on case data. The technical integrity of the research is the primary priority of each paradigm that follows. Nevertheless, Divulge seeks aid investigator in overcoming the adversarial and technological difficulty by supporting the technical.

In conclusion, the only framework for digital forensics that shared its objectives and approach was disclose. More exact all of the methods and the consideration aimed to support The invigilation and increase one of the two ways by offering a set of guidelines, a logic framework, or a device for automated correlations.

Contribution 1: offering guidelines that follow and support the inquiry process step-by-step, DISCLOSE is the first similar effort to do so. Possible attack strategies will be examined.
Prior works offer the user generic recommendations/guidelines without adapting them or the stage of the inquiry.

Contribution 2: DISCLOSE customises future suggestions depending on the results by taking into account the investigator's remarks at each level of the study.

Because (a) many strike activities lack the use of a frailty, (b) many attack, breach[7] actions include more than one weakness, and (c) a frailty is only able to capture the system's weaknesses rather than adversarial behaviour and cultural knowledge, all prior rationalisation frameworks only deem an individual instance of an attack to be a weakness, which is not indicative of the existing threat[6] terrain..

Contribution 3: By employing adversarial TTPs that are consistent with the present danger environment, DISCLOSE improves this by enabling representational modelling of the investigation and enhancing its efficacy. The framework's capacity for cyber kill[12] and social kill[14] in a broader variety of situations. A TTP[1] can express in more detail and is more complex. hostile behaviour as well as its aim and method of execution. As a result, disclose can employ TTPs to circumvent issues like zero-day vulnerabilities. Despite the fact that some works used the mitre att&ck[1] TTPs, comparable to disclose[1], they did not make use of any additional knowledge base components, such as the characteristics.

Contribution 4: By extracting several factors from the att&ck stix[1] repository and utilising those variables to determine the probabilistic relations between TTPs, disclose[1] improves on this by giving the investigator data-driven decision assistance.

Contribution 5: Last but not least, in contrast to previous studies, DISCLOSE employs a well-known attack[4] life cycle model[11] in addition to some of the variables obtained from the att&ck stix repository. As a result, our framework is not just data-driven but also conscious of an event's underlying structure, enabling it to provide suggestions that are even better.

III. PROPOSED SYSTEM

The system is a hybrid model making use of two encryption algorithms while transferring data from one module to another. Data will be analyzed to optimize the process. From the report, the analysis process[17] provides the decision[3] making solution. Cyber forensic contains steps to investigate or collect the data It is defined as the processes and tools used in investigations and gathering evidence[13]. Some of the instructions will be provided as a default such as category wise. By analyzing the investigation report, the process will be optimized to reduce the investigation process. In our system, we make use of AES Algorithm and blowfish Algorithm which enhances the security[5] of the information as well as make it difficult to decrypt. Two different algorithms (i.e) AES Algorithm and Blowfish algorithm can be used for efficient data functionality when compared to disclosure framework technology. The AES algorithm can be used for storing large amounts of information and provides enhanced data security. An alternative to the DES The symmetric keying method is the cryptosystem known as blowfish. A symmetric, 64-bit block encryption with adjustable length is called blowfish. To replace the outmoded Data[1] Encryption Standard (DES) and IDEA encryption techniques, or international data encryption algorithms, Bruce Schneier developed a "general-purpose algorithm" in 1993. Blowfish is unpatented, substantially quicker than DES and IDEA, and freely accessible for all purposes. Unfortunately, because of its short block size, which is seen as unsafe, it couldn't totally replace DES. This approach is put into practise where the general public has complained using the created web application. This technique is being used at this location withthe intention of encrypting any data provided by the public. The Rijndael technique, often known as the AES cryptosystem, is a symmetric block encryption method that uses blocks and chunks with a size of 128 bits. It then arranges all of these blocks into the ciphertext after separately encrypting each one. Its base is an SP network, often known as a subtraction network. It includes a variety of core elements, some of which need bit shuffles and others of which require the replacement of specified entries with particular outcomes (permutations). Information is transmitted through Many processes are required to comprehend how AES works. If a block consists of block size, with one byte stored in each cell, a 4x4 matrix conveys the information within one block. Key expansion is the process of extending the unique identifier into (n+1) keys, where n represents the total number of cryptography rounds. As a result, 16 cycles are required to create a 128-bit key, yielding a total of 11 keys (10+1).

ALGORITHM IMPLEMENTATION:

1. BLOWFISH ALGORITHM:

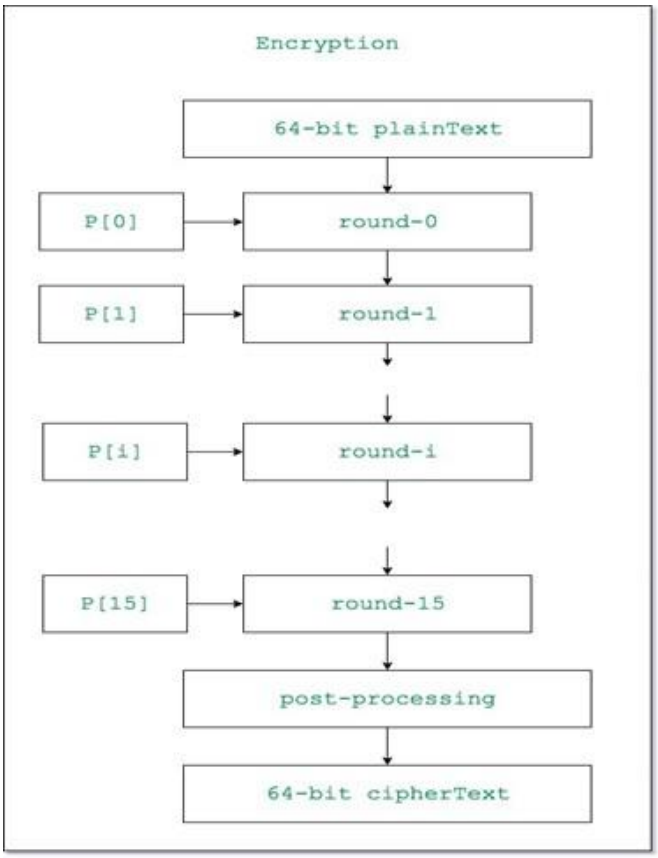


Fig 1. Blowfish Algorithm

The two primary parts of the blowfish algorithm are as follows:

Encryption process is one. A 16-round Feistel network is used to encrypt data; each round includes a crucial permutation and a crucial replacement.

To encrypt data in Blowfish, the replacement strategy is combined with big, key-dependent S-boxes. Every encryption method uses XORs, a type of digital circuit, and 32-bit word extensions are often used.

Key expansion and subkeys are two. The Blowfish algorithm, which makes extensive use of subkeys, depends on them. Before any encryption or decryption can occur, these subkeys are pre-calculated.

Example of the Blowfish encryption and decryption technique:

Suppose Blowfish will be used to encrypt the message "Hello world." The procedures are as follows:

1. The initial input, "Hello world," consists of seven letters plus a space, for a total of eight bytes (64 bits).

2. A 32-bit division of the input is made. Key expansion produces a value termed P1, which is then XORed with the remaining 32 bits, "Hi w," to produce the desired result. (Note: P stands for a prime number, which can only be divided by itself and by one other.)

3. Then, a transformational F-function (F In) runs through P1, dividing the 32 bits into 4 bytes each and passing them to the four S-boxes.

4. The third value from the third S-box is XORed with the first two values from the first two S-boxes.

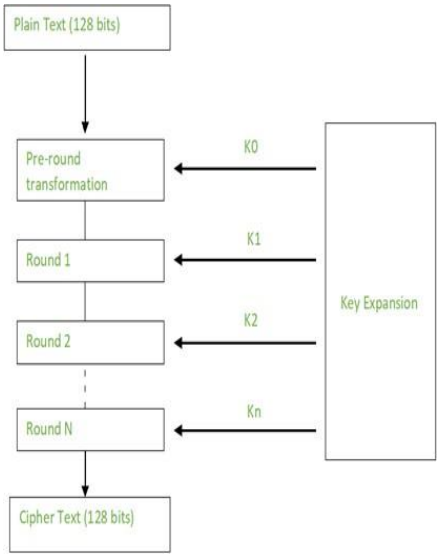
5. To create 32 bits as the output, this result is added to the fourth S-output. box's

6. The result of F In is XORed with the right 32 bits of the input message "orld" to produce output F1'.

7. Then, F1' and P1' are used to swap out the left and right halves of the message, respectively.

8. There are a total of 16 rounds in which the same procedure is performed for each new P-array member.

9. The outputs P16' and F16' are finally XORed with the P-final array's two entries, P17 and P18, after 16 rounds. These are then recombined to create the input message's 64-bit ciphertext.

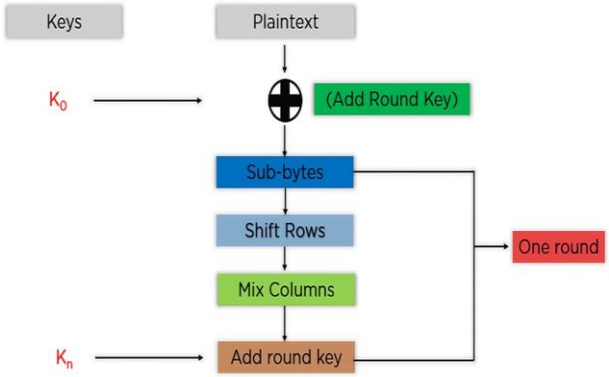


AES Algorithm

2. AES ALGORITHM

To understand how AES works, you must first understand how data is transferred between different phases. Given that a A single block has 16 bytes, with one byte kept in each unit, and is carried through a dimensional array.

State vectors can be used to explain the matrix in the prior example. Similar operations are performed on the by (n+1) keys, where n is the number of crypto rounds, using the originating key. As a result, 16 cycles are necessary to create a 128-bit key, producing a total of 11 keys (10+1).



1. The bytes are switched out:

Initially, the block text's bytes are modified in according to the directions provided by the specified S-boxes (short for substitution boxes).

2. Rearranging the rows

The matrix's four rows are all moved to the left.
On the right side of the row, any entries that "drop off" are reattached. The continuous stream is suitable for shift.

3. The top row is still existent

4. The second row is moved left one (byte) position.

5. The third row has moved two positions to the left.

6. The fourth row has been moved three positions to the left.

7. The upshot is a new matrix with the same 16 bytes, but withalteredrelative-positions.

a. Conflating the column:

Each four-byte column is now altered using a special mathematical procedure. This method takes four bytes, one for each column, and returns four brand-new bytes, one for each column, to replace the original four bytes. An additional matrix with 16 additional bytes is the result. Please be aware that this step is not part of the final round.

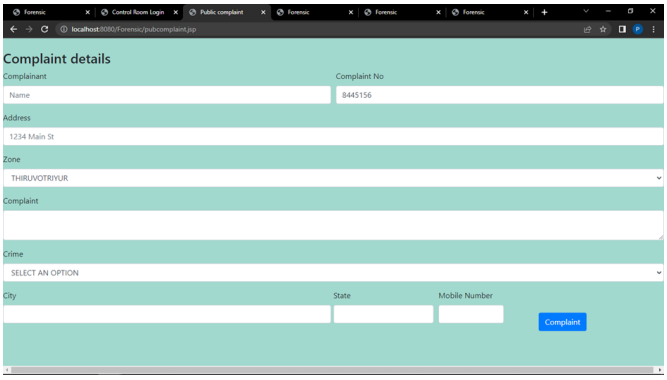
b. Adding the round key:

The matrix's 16 bytes—now regarded as equivalent to 128 bits—are XORed with the round key's 128 bits. In the case where this is the last round, the output is the ciphertext. If not, the procedure is repeated, translating the output 128 bits into 16 bytes.

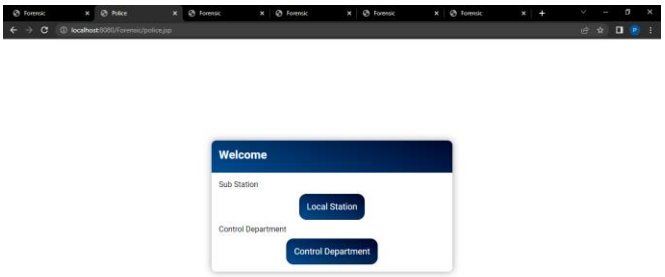
IV. RESULTS AND DISCUSSIONS

HOME PAGE OF WEB APPLICATION

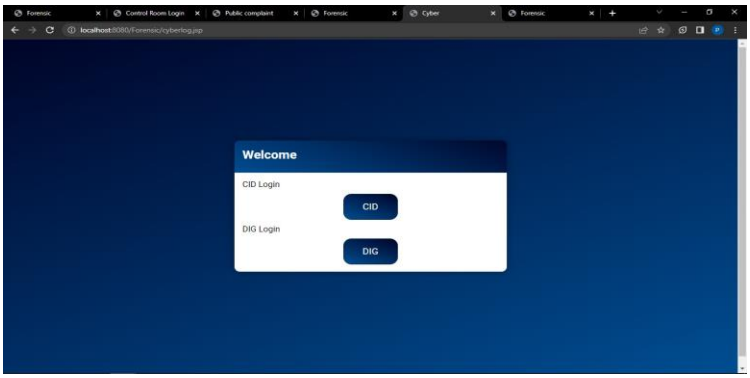
PUBLIC COMPLAINT PAGE



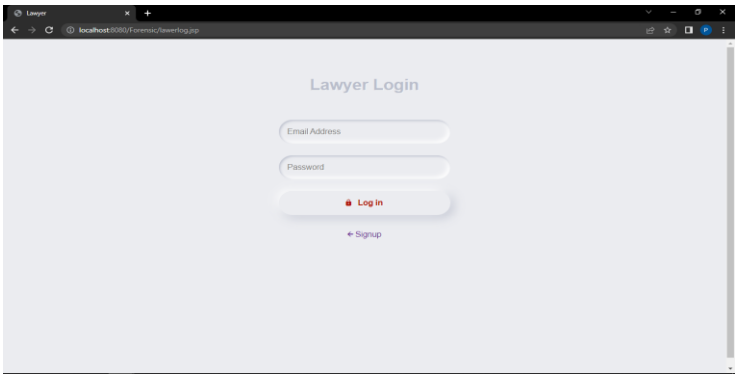
POLICE LOGIN PAGE



CID & DIG LOGIN PAGE



LAWYER LOGIN PAGE



V. CONCLUSION

Digital forensics involves several processes, including identification, collection, acquisition, and preservation. examining and presenting technological proof. Before it can be presented in court as evidence, digital evidence needs to be verified. The case and its elements ultimately decide the evidence artefacts and evidential methods utilised for static or active gathering. Genuine evidence needs to be

material, competent, and authenticated. The development of this software used a modular approach. Every module in this system has undergone successful testing using accurate data.

The model for the upcoming project was created with two experts or administrators to review it and provide suggestions to the public with permission services. A first phase will be deployed for a variety of complaint categories, including homicide and kidnapping, robbery, etc. One more stage of cybercrime

VI. REFERENCES

[1] Antonia Nisioti; George Loukas; Aron Laszka; Emmanouil Panaousis, "Data-Driven Decision Support for Optimizing Cyber Forensic Investigations", IEEE, vol. 16, January 2021

[2] Seonghyeon Gong; Changhoon Lee, "Cyber Threat Intelligence Framework for Incident Response in an Energy Cloud Platform", Seoul National University of Science and Technology, vol. 10, January 2021

[3] Antonia Nisioti, George Loukas, Stefan Rass, Emmanouil Panaousis, "Game Theoretic Decision Support for Cyber Forensic Investigation", University of Greenwich, London, UK, vol. 21, August 2021

[4] Zhun Zhang; Qihe Liu; Shilin Qiu; Shijie Zhou; Cheng Zhang, "Unknown Attack Detection Based on Zero-Shot Learning", IEEE, vol. 8, October 2020

[5] K. Finnerty, S. Fullick, H. Motha, J. N. Shah, M. Button, and V. Wang, "Cyber security breaches survey 2019," Dept. Digit., Culture, Media Sport, London, U.K., Tech. Rep., Apr. 2019

[6] V. Diaz, D. Emm, and C. Raiu, "Kaspersky security bulletin 2019: Advanced threat predictions for 2020," Kaspersky Lab., Moscow, Russia, Tech. Rep., 2019

[7] *Cost of a Data Breach Report 2019*, IBMSecurity, New York, NY, USA, 2019

[8] G. Horsman, "Formalising investigative decision making in digital forensics: Proposing the digital evidence reporting and decision support(DERDS) framework," Digital Investigation, vol. 28, pp. 146–151, 2019

[9]J. Navarro, A. Deruyver, and P. Parrend, "A systematic survey on multistep attack detection," *Comput.Secur.*, vol. 76, pp. 214–249, Jul. 2018

[10] V. S. Harichandran, F. Breitinger, I. Baggili, and A. Marrington, "A cyber forensics needs analysis survey: Revisiting the domain's needs a decade later," *Comput. Secur.*, vol. 57, pp. 1–13, Mar. 2016

[11] Rabail Shafique Satti and Fakeeha Jafari, "Domain Specific Cyber Forensic Investigation Process Model", Journal of Advances in Computer Network, vol. 3, March 2015

[12] L. Martin. (2014). *Cyber Kill Chain*.[Online]. Available: http://cyber.lockheedmartin.com/hubfs/GainingAdvantageCyberKillChain.pdf

[13] J. Williams, "Acpo good practice guide for digital evidence," Metrop. Police Service, Assoc. Chief Police Officers, GB, London, U.K., Tech. Rep., Mar. 2012

[14] Humaira Arshad, Saima Abdullah, Moatsum Alawida Abdulatif, "A Multi-layer Semantic Approach for Digital Forensics Automation for Online Social Networks", February 2012

[15] S. Barnum, "Standardizing cyber threat intelligence information with the structured threat information expression (STIX)," *Mitre Corp.*, vol. 11, pp. 1–22, Jan. 2012

[16] L. F. da Cruz Nassif and E. R. Hruschka, "Document clustering for forensic analysis: An approach for improving computer inspection,"IEEE transactions on information forensics and security, vol. 8, no. 1,pp. 46–54, 2012

[17] S. Alharbi, J. Weber-Jahnke, and I. Traore, "The proactive and reactive digital forensics investigation process: A systematic literature review," in Proc. Int. Conf. Berlin, Germany: Springer, October 2011

[18] SunHo Cho, Hyuk-Chul Kwon, "Cyber Forensics Ontology for Cyber Criminal Investigation" Digit. Invest., vol. 3, January 2009

[19] A. Brinson, A. Robinson, and M. Rogers, "A cyber forensics ontology: Creating a new approach to studying cyber forensics," *Digit. Invest.*, vol. 3, pp. 37–43, Sep. 2006

[20] K. Kent, S. Chevalier, T. Grance, and H. Dang, "Guide to integrating forensic techniques into incident response," *NIST Special Publication*, vol. 10, no. 14, pp. 800–886, 2006