iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** II **Month of publication:** February 2018

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor : 6.887
Volume 6 Issue II, February 2018- Available at www.ijraset.com

Digital Forensics

Pooja Nagdev[1], Sahil Jagiasi [2], Ekta Chawla3, Jai Kukreja[4]
[1]*Assistant Professor, Computer Engineering, V.E.S.I.T. Mumbai, India*
[2. 3. 4.]*Computer Engineering, V.E.S.I.T. Mumbai, India*

Abstract*: *The recent development in Information Communication Technology (ICT) has made changes in every aspect of our Life. These changes are taking us towards the dream of "DIGITAL INDIA". The positive influence of Digital world on Knowledge, trade and business and Communication is no doubt remarkable. However, the dark side of it deteriorates its peaceful usage that is Digital Crimes .Digital Crimes are defined as any illegal activities practiced by or done via digital device. Unlike "traditional "crimes Digital crimes present a real dilemma due to the fact that criminals' identity may be hidden. Digital Forensics along with the process of finding the digital evidence and tools used in digital forensics.
Keywords: Information Communication Technology (ICT), Digital Crimes (DC), Data Analysis (DA)

I. INTRODUCTION

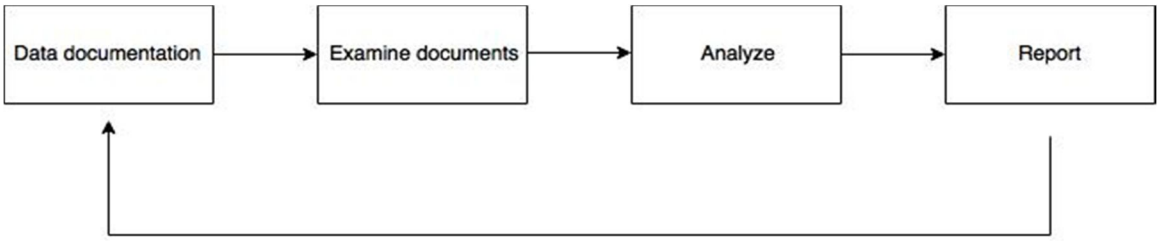
Computer forensics is the application of investigation and analysis techniques to gather and preserve evidence from a particular computing device in a way that is suitable for presentation in a court of law. The aim of computer forensics is to perform a systematic investigation during maintaining a documented chain of evidence to catch in exactly what happened on a computing device and who was responsible for it.

Forensic investigators typically follow a standard set of procedures: After physically isolating the device in question to make sure it cannot be accidentally contaminated, investigators make a digital copy of the device's storage media. Forensic scientists collect, preserve, and analyze scientific evidence during the course of an investigation. While some forensic scientists travel to the scene to collect the evidence

themselves, others occupy laboratory role, performing analysis on objects brought to them by other individuals.

II. METHODOLOGY

A text message corpus has been developed and basic experiments were conducted in order to show information about the corpus and to demonstrate natural language processing (NLP) principles and machine classification based on supervised learning algorithms. Applicability and limitations of the corpus are discussed. A simple methodology for extracting features from the corpus is proposed.



A. Mobile Forensics

Mobile device forensics is a sub-branch of digital forensics relating to recovery of digital evidence or data from a mobile device. It differs from Computer forensics in that a mobile device will have an inbuilt communication system (e.g. GSM) and, usually, proprietary storage mechanisms. Investigations usually focus on simple data such as call data and communications (SMS/Email) rather than in-depth recovery of deleted data.

B. Network Forensics

Network forensics is concerned with the monitoring and analysis of computer network traffic, both local and WAN/internet, for the purposes of information gathering, evidence collection, or intrusion detection. Traffic is usually intercepted at the packet level, and either stored for later analysis or filtered in real-time. Unlike other areas of digital forensics network data is often volatile and rarely logged, making the discipline often reactionary.

C. Need for forensics

Theft of Intellectual Property.

1) Financial frauds.

2) Hacker system Penetration.

3) Distribution and execution of viruses and worms.

III. PROCESS OF DIGITAL FORENSICS

The process of digital forensics involves many steps. The first step is to find out whether the information is sufficient or not for the analysis process. The flow chart displays the complete process if the data is completely available then the analysis phase is done with the tools to generate the report. The analysis will include the most active users in the social media and the data of the most activated users is accomplished. The examination of the data includes the vulnerability involved in the data. Basically it does the sentiment analysis on the given data. It results into the positive or the negative drug involved in the activity. The paper defines to do the stylometry involved in the process. The source of the data being sent and the ip address of the related network from which the data is sent. The report is being generated regarding the analysis performed and submitted in the court of law for evidence of crime being suspected. Identification is started after the analysis. Collect information: The data is collected from the social media such as whatsapp, twitter or facebook. Authenticate information: The collected data is authenticated and further provided for analysis . Analysis of the information: The semantic analysis is accomplished. Report: The report is generated and provided to the user.

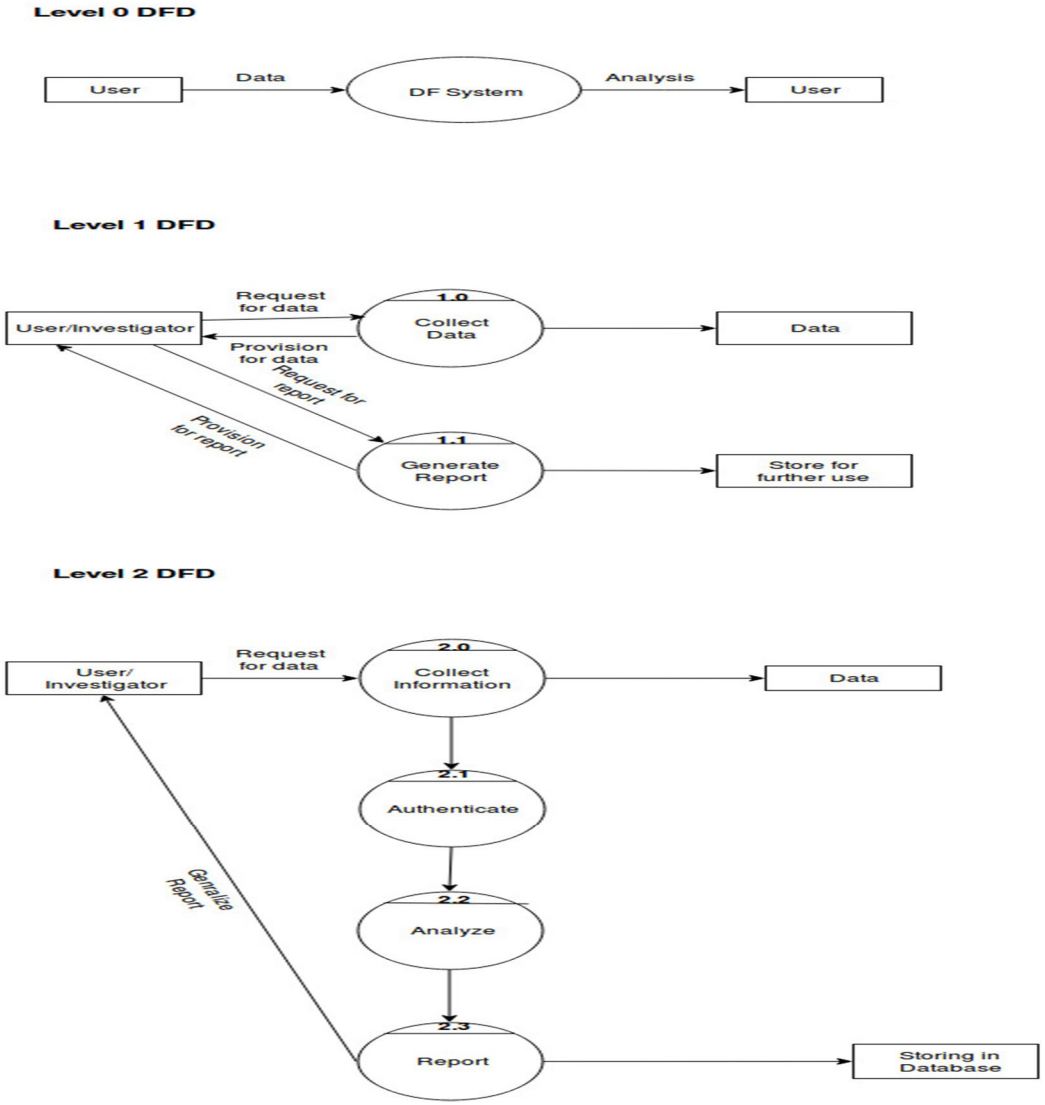


Figure: DFD of DIGITAL FORENSICS

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue II, February 2018- Available at www.ijraset.com

IV. STEPWISE MEDIA ANALYSIS

A. *Basic Steps*

1) Copy of Image: Acquitisition of data of copied image taken from the device.
2) Perform Hash Calculation of the data.
3) Conducting recovery procedure to retrieve.
4) Scan all the deleted files.

B. *Analysis of Hidden evidence*

The sensitive data can be hidden by the accused in file system such as stack.

1) Hard disk
2) File system tables
3) File deletion

The process of digital forensics involves many steps. The first step is to find out whether the information is sufficient or not for the analysis process. The flow chart displays the complete process if the data is completely available then the analysis phase is done with the tools generate the report.

1) The analysis will include the most active users in the social media and the data of the most activated users is accomplished. The examination of the data includes the vulnerability involved in the data. Basically it does the sentiment analysis on the given data. It results into the positive or the negative drug involved in the activity. To perform sentiment analysis on the given data on social media .
2) To determine it is a positive or a negative drug.
3) Conceptual use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and positive and the negative information.
4) Retrieve the ip address and the network from which the data is being sent.

The report is being generated regarding the analysis performed and submitted in the court of law for evidence of crime being suspected. Identification is started after the analysis.

C. *Log file analysis*

During investigation to recognize malicious activities by mining log files. Access logs can contain vast amount of data regarding user activities

Analysis steps

1) Input a server log file.
2) Identify each sessions.
3) Using the search algorithm find the required data.

V. PROPOSED SYSTEM

The proposed solution caters to all the requirements of the user and overcomes all the deficiencies in the existing system and as well have all the functional requirements implemented with intelligence of its own.

Algorithms used are decision tree : The algorithms used in the system are Decision tree. A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. Divide and conquer method is adopted.

Target concept *Interesting*? : *Document* $\rightarrow \{+, -\}$

1. Represent each document by vector of words
   - one attribute per word position in document
2. Learning: Use training examples to estimate
   - $P(+)$
   - $P(-)$
   - $P(doc|+)$
   - $P(doc|-)$

Naive Bayes conditional independence assumption

$$P(doc|v_j) = \prod_{i=1}^{length(doc)} P(a_i = w_k|v_j)$$

VI. CONCLUSION

We Indians are on the way to achieve the dream of Digital India!. Digital Crimes, on the other side of the coin, are increasing on the faster pace. Thus, Boosting the need of Digital Forensics. It has evolved a lot with time and hence the DF tools also. To make the process effective the discipline is subdivided into various branches where specialized tools are available for a particular branch. As every Forensic tool is associated with some or other limitation proper tool should be used as per the requirement of case in hand.

Though tools provide the data from the suspects' device it is still difficult for the investigators to analyze the complete data and produce the required evidence in court of Law. The problem is proportional as the number of devices involved in the crime increases. The application of Knowledge Management classification algorithms to Mobile Forensics helps to isolate potential evidence from mobile phones or any other digital device with both internal memory and communication ability.

Natural language processing (NLP) and machine classification as discussed were applied to mobile device forensic analysis .Feature extraction by analyzing the linguistic patterns would be helpful to classify text messages as related to crime or not. Further research recommendations include determination of the frequency of text messaging between criminal suspects versus the general population, calculating the average time span between sent and received messages in text message conversation threads, and the analysis of additional types of crime and suspect behavior.

VII. FUTURE WORK

A. To perform sentiment analysis on the given data on social media .

B. To determine it is a positive or a negative drug.

C. Conceptual use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and positive and the negative information.

D. Retrieve the ip address and the network from which the data is being sent.

REFERENCES

[1] A quantitative approach to Triaging in Mobile Forensics by Fabio Marturana, Gianluigi Me, Rosamaria Bertè, Simone Tacconi, 2011 International Joint Conference of IEEE TrustCom-11/IEEE ICESS-11/FCST-11.

[2] Text Message Corpus: Applying Natural Language Processing To Mobile Device Forensics Daniel R. O'Day And Ricardo A. Calix, Multimedia And Expo Workshops (ICMEW), 2013 IEEE International Conference

[3] Cyber Forensics by LabSystem (I) pvt. Ltd.

[4] Mobile Device Forensics: Extracting and Analysing Data from an Android-based Smartphone Normaziah A. Aziz, Fakhrulrazi Mokti, Mohd Nadhar M. Nozri, 2015 Fourth International Conference on Cyber Security, Cyber Warfare, and Digital Forensi

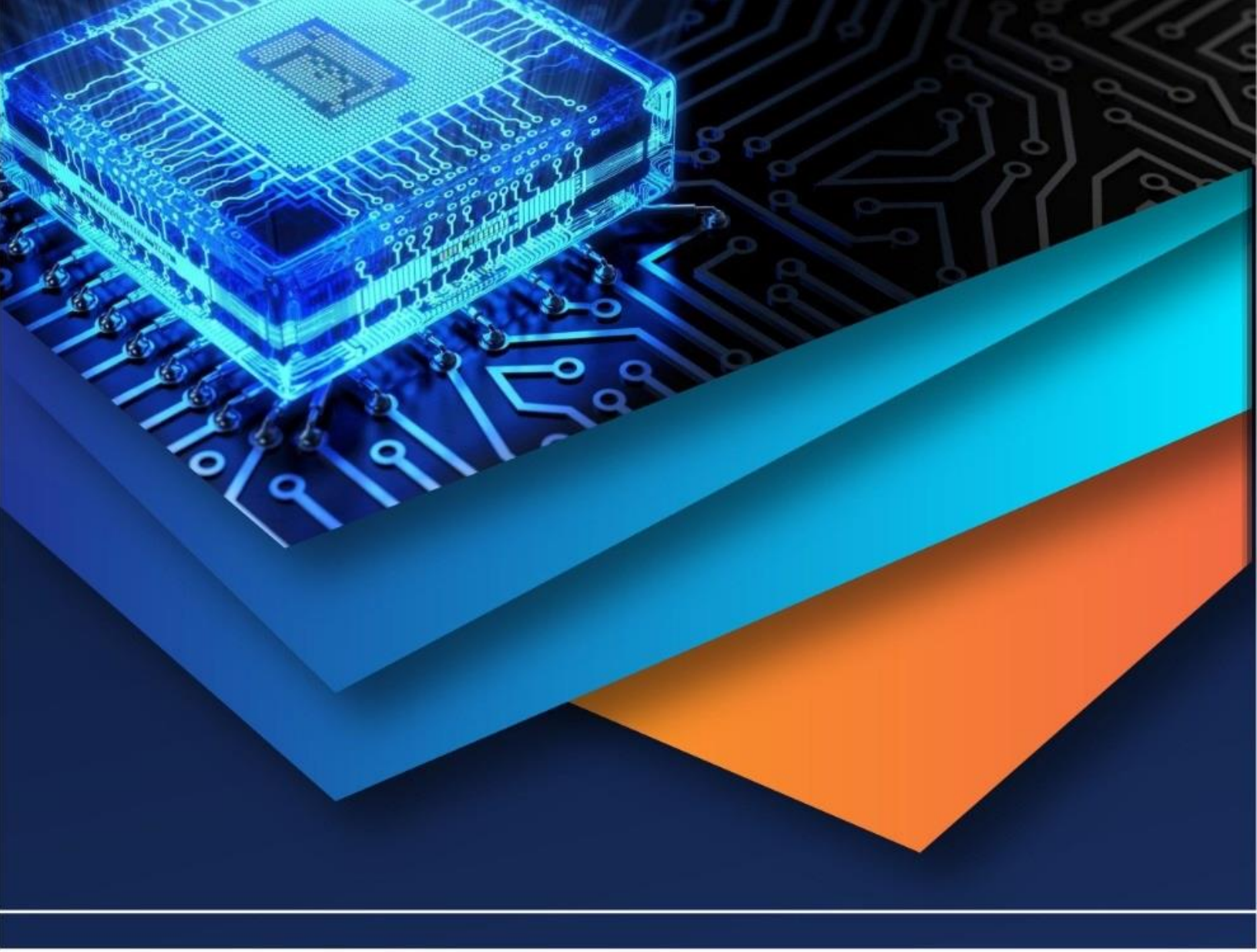
[5] Smartphone Forensics Analysis: A Case Study, Mubarak Al-Hadadi and Ali AlShidhani, International Journal of Computer and Electrical Engineering, Vol. 5, No. 6, December 2013

[6] https://en.wikipedia.org/wiki/Digital_forensics

[7] https://en.wikipedia.org/wiki/forensics.

[8] https://en.wikipedia.org/wiki/MOBILedit

[9] https://en.wikipedia.org/wiki/List_of_digital_forensics_tools



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY