International Journal for Electronic
Crime Investigation
(IJECI)

Digital Forensics Research and Services Center
Lahore Garrison University Lahore, Pakistan.

SCOPE OF THE JOURNAL

The IJECI is an innovative forum for researchers, scientists and engineers in all domains of computer science and technology to publish high quality, refereed papers. The journal offers articles, survey and review from experts in the field, enhancing insight and understanding of the current trends and state of the art modern technology. Coverage of the journal includes algorithm and computational complexity, distributed and grid computing, computer architecture and high performance, data communication and networks, pattern recognition and image processing, artificial intelligence, cloud computing, VHDL along with emerging domains like Quantum Computing, IoT, Data Sciences, Cognitive Sciences, Vehicular Automation. Subjective regime is not limited to aforementioned areas; Journal policy is to welcome emerging research trends in the general domain of computer science and technology.

SUBMISSION OF ARTICLES

We invite articles with high quality research for publication in all areas of engineering, science and technology. All the manuscripts submitted for publication are first peer reviewed to make sure they are original, relevant and readable. Manuscripts should be submitted via email only.

To submit manuscripts by email with attach file is strongly encouraged, provided that the text, tables, and figures are included in a single Microsoft Word/Pdf file. Submission guidelines along with official format is available on the following link; www.research.lgu.edu.pk

Contact: For all inquiries, regarding call for papers, submission of research articles and correspondence,kindly contact at this address:
IJECI, Sector C, DHA Phase-VI Lahore, Pakistan
Phone: +92- 042-37181823
IJECI@lgu.edu.pk

CONTENTS

Bank Frauds Using Digital Devices and the Role of Business Ethics

Aftab Ahmad Malik[1], Mujtaba Asad[2], Waqar Azeem[3]

[1]Professor Department of Computer Science, Lahore Garrison University (LGU), Pakistan

[2]PhD Scholar, School of Electronics Information & Electrical Engineering,Shanghai Jiao Tong University Shanghai China

[3]Senior Lecturer, Department of Computer Science Lahore Garrison University (LGU), Pakistan

dr_aftab_malik@yahoo.com[1], asadmujtaba@sjtu.edu.com[2] , waqar.azeem@lgu.edu.pk[3]

Abstract

The paper focuses on the use of digital devices to commit the banking frauds in collaboration with bank employees. There have been enormous cases of small and large size embezzlements, illegal deductions and frauds with customers. Most of these frauds are accomplished using digital devices. The recent developments in computer networking, detection of illegal access to accounts by unauthorised attackers and hackers have made the task of protection to customers easier. This paper considers the role of business ethics in banking sector which may serve as a deterrent force for the employees. Several studies have revealed that the implementation of code business ethics produced encouraging results.

Keywords: Digital Banking, White Collar Frauds, Business Ethics

1. Introduction

In the recent days due to the rapid increase in business activities the role of banking sector has become challenging. The banking business mostly is carried out using computerized systems and automated knowledge based software and expert systems to operate and manage the data warehouse, data marts, databases of customers and firms. According to [21], the customers do have fear of frauds in the transaction with the banks. The customer demands fast processing of the transactions. Various I.T. systems provide the bank employees to commit frauds with the customers. This demands utmost honesty by the operating bank officers. According to [21], the digital bank frauds also create bad reputation of the bank and imply inefficient operational behaviour in spite of high speed processing facilities and technical assistance. Therefore, the business ethics has to play a pivotal role in banking sector. References [1] to [19] shall be basis code of ethics presented in this paper.

2. Typical Bank frauds

Due to digitalisation in banking systems remarkable increase in offences and frauds are being committed both at large and small scale. The common Bank frauds are related to credit cards, Money Laundering, Internal misappropriations, mobile banking, through false Emails and misrepresentation of customers.

3. Business Ethics

Behavioural assumptions guide the economic theories and policies. The economic models are constructed and tested on the basis of behaviour; therefore, ethics, values and economic development go side by side. Values embrace the entire spectrum of subjective attitudes, judgments, interpretations and behavioural preferences.

Truthfulness and trustworthiness in statements and actions constitute the basis of business ethics. Continuing professional development needs enlargement of knowledge and research. The most salient features of business ethics are:

(i) Loyalty to client, employer and institution;

(ii) Avoidance of conflict of interests, bribery, harm to outsiders and public;

(iii) Respect of autonomy, privacy and of trade secrets;

(iv) Non-exploitation of Bank clients and fairness;

(v) Proper treatment of subordinates;

(vi) Protecting honor and dignity of profession.

4. Operationally Feasible System for Business Ethics

In order to establish an operationally feasible system for business ethics, let us first analyze the notions of 'good', 'right' and 'duty'. The ultimate question of moral philosophy revolves round these important factors. Then the questions arises, how the pleasure should be exercised? Should it he exercised in a fair, impartial, judicious, honorable and in a moral manner without bias, victimization or favoritism?

Rationalistic ethics must treat virtue or character as itself part of the good to be promoted. The moral judgment in business is thus judgment of value. Morality is a categorical imperative in recognizing that performance of duty is the highest good. The moral judgements are normally accompanied by emotion.

5. Maxims of Morality

The ethical maxims such as prudence, benevolence and equity are self-explanatory, though they do not resemble mathematical axioms and theorems, but they are applicable (like mathematical formulae) to human conduct. Similarly, the factors like aesthetic judgements for objective validity and impartiality, good faith (in legal terms), humanity and social affection create a temperament in the field of business ethics.

Other ideal values such as value of knowledge, culture, intellectual activity, veracity, love of truth, purity, temperance and humility also do have a great deal of positive impact on practical ethics.

6. The Ethical Conduct in Business

The ethical conduct requires that one be free and accepts the responsibility that freedom bestows. A distinction can be drawn between ethical, legal and professional conduct. Ethical conduct relates to what is right and wrong in respect of voluntary behaviour. The usual norms of business and standards of conduct that are normally expected in business, so that the way the professionals conduct their activities and profession can be explained, judged or interpreted. Whereas, banking law, pattern of banking regulations, statutes, rules, legal standards and international conventions need to be strictly followed.

Apart from legal and professional standards, the spiritual, political and social values do have moral impact on the formation of structure of the business ethics. For example, restlessness adverse law and order situation and irresponsible attitudes of people in a country directly influence and adversely affect the. Economic values, investments and productivity.

7. The Business Ethics of Islam

This derives wisdom from references from [1] to [19]. The binding force behind Business Ethics in Islam is faith and fear of Allah the Almighty. "Every religion has distinctive virtue and the distinctive virtue of Islam is modesty".' In Islam "Every good work is charity"(Hadith Mubarik), where the term 'good' means what is acknowledged to be good, fit, decent, becoming, good conduct, obligingness, kindness, legitimate favor, bounty, and obedience towards Allah. HazratAmru Abbas" (91,2210}; asked the prophet (PBUH) "What is Islam?" Prophet Muhammad Peace be upon him said, "Purity of speech and hospitality". Islam changed the course of world history. The life of Prophet Muhammad Peace be upon him is the only perfect and ideal model of humanity, which helps to be enlightened in every aspect, department and field of existence as a legislator,

economist, statesman, preceptor, judge, trader, philosopher, teacher, missionary, intellectual, general, soldier, patience, selflessness and sacrifice for others.

The Prophet of Islam (PBUH) is the only perfect model of a man and leader comprehending all virtues and a greater man ever existed in the annals of the world. No such personality is to exist in future for refinement of morality and humanity like him. The prophet of Islam (PBUH) is a model of excellence to follow. "He is sent down as mercy for all the worlds", a model of mercy, self-control and truth.

The Prophet of Islam (PBUH) never preferred his personal wants to the needs of his family. He came of a family exalted in prestige and honor in the whole of Arabia. Due to his sincerity and truthfulness, he took charge of HazratKhudijatulKub'ra's business and made great profit by his honesty, intelligence and added to her fortunes. She also fully appreciated the noble, genius and commanding personality of Mohammad (PBUH).

The basic ethical principles stated in Qur'an are actually ethical judgements dealing with the nature of business morality and the spirit how it is to be practiced. The spirit behind moral ethico-religious dynamics of Islam is so bread that it is practically boundless. The basic principles of Qur'an and Sunnah are to be applied in settling our practical problem of organizing the life of the community of believers. Qur'an is a perpetual source of guidance.

Islam possesses an ideal economic system based on the Heavenly virtue, rules of goodness, fair play, justice, morality and business ethics. "Holy Quran" as a whole is a reply to our queries, worries and miseries". As a dynamic system and safe anchor for humanity, Islam embodies virtues necessary for the individual and the community. Its unique distinction and greatness lies in its emphasis on practice.

Islam advocates loyalty, heedfulness, truthfulness, patience, discipline, steadfastness, courage, endurance, kindness, tolerance in speech and action, generosity, forgiveness, virtue, reliance on Allah, protocol with public business and straightforwardness.

Islam prohibits falsehood, hypocrisy, deception, lavish life style, conceit and vanity, exploitation, gossip and rumour, intrigues, greed and bribe, illegal gratification, prodigality and treachery. In Qur'an Allah says about those of His men whom neither business nor trading distracts from remembering him. Muslims must conduct their affairs by mutual consent or consultation.

It is clear from the above that Islam is the seat of justice and faith upon earth which teaches furtherance of good and repression of evil. The whole corpus of Qur'an, Hadith and the holy life of' Mohammad (PBUH) constitutes the Islamic Ethics, which implies the correct performance of duties and the right understanding of moral life and its doctrine.

The essence of 'Akhlak' is considered together with politics and economics as a part of practical philosophy. Therefore, it is the science of virtues and the way how-to acquires them, of vices, and the way how to guard against them. Wisdom, valnor, temperance and justice of course are associated with the Islamic 'Akhlaq' along with a large number of subordinate virtues of operational value.

In this section, we have highlighted the ethical aspects needed in banking business, implicitly; whereas, the holy injunctions of sha'ri'ah have been explicitly portrayed. In the next sections, explicit discussion follows with special emphasis and application to banking sector.

8. Identification of Unethical Behaviour in Banking Sector

Negligent behaviour, fraudulent behaviour, deviant behaviour and moral dilemmas are common in banking sector. Unethical norms in various functions of banking are observed, such as receiving, collecting, transferring, and paying, lending, dealing, investing, exchanging and servicing money and money claims both domestically and internationally.

Private Banks, industrial, commercial, multistate, multipurpose banks and holding companies are often seen practicing unethical conduct in one way or the other. The private commercial banks working in credit to business and operations in the fields of "thrifts", savings and time deposits, mortgages, lending and dealing with consumer loans, personal loans and securities business also have been involved in well-known unethical practices. Commercial banks and the "thrifts" are affected by new products and services of development by securities firms including money market funds.

The trust companies and financial enterprises are also engaged in many kinds of illegal and unethical practices while transferring money at home and abroad, collection and exchanges while providing services as stock transfer agent at the time of bringing the buyer and the seller together as travel agent.

8.1 Disinformation about Trade Data

Sometimes incorrect 'trade data' is deliberately floated to deceive and cause damage to other business counterparts by the businessmen. The indices of industry and commerce, statistical indicators of business conditions, the comparative parameters of business statistics on which business volume, activity, credit supply, price trends, profit prospects and investment opportunities depend are prepared carelessly due to negligence, inexperience or deliberately with dishonest intent.

8.2 Role of Auditors

Certain other important limitations and precautions which are to be observed in the profession are not followed properly. Substantial and serious errors are committed in the banking business due to poor accounting method and missing information. This leads to opportunities for frauds. Frauds also arise out of false declarations. Goods in transit may also be a source of fraud by certain companies. Some bank frauds are committed in collaboration with the auditors.

9. Political Influence and Accountability for Unethical Practices in Banking

Fraudulent extraction of wealth and money laundering are unethical, very serious and also illegal acts. There are also matters of financial bungling through banks. Tales of plundering the national exchequer are not uncommon. However, in the absence of any code of ethics, we will never know the whole truth. Financial bungling through banking sector and everything in this regard has been running as a desirable fashion. A notable flaw in such situations has been due to the absence of a clear demarcation between the 'practice of polities" and 'business interests'. Hence, as per democratic traditions, there is a 'need for clearly formulated rules aimed at eliminating the use of political power for the furtherance of politician's business interests by the banking sector.

9.1 Adverse effects on Banking Sector

Foreign investments, the world affairs,

international money, world trading and economic crisis also give rise to many, unethical policies, which adversely affect banking business. The banking business collapses because banks and government lent far more money irrationally than they had when debtors defaulted. When investments are dried up, they lead to unemployment and protectionism.

9.2 Accountability

The process of accountability is the most important feature of democratic order. There can be no sensible argument against any government starting such a process. Accountability can only be maintained if it is fair and free from favoritism or victimization. The process of accountability should cover all types of financial wrong doings in the country.

The process of accountability may be developed by removing the mistrust surrounding and bringing in the business ethics and moral conduct of all concerned. The businessmen extract heavy loans, from banks and do not return them. Many bungling acts are performed by account holders by means of hidden companies and the money is extracted from the banks.

10. Essence of Operational Code of Ethics in Banking Sector

In banking business, three kinds of standards are relevant to professional life:
·   Technical,
·   Economic and
·   Moral.

Should the code of ethics be restricted to moral norms of the business? The moral community is composed of all those beings towards which we have moral obligations. To make a code really effective, feasible and operational, one must get entire profession to agree on its foundation.

Every genuine profession has a code of ethics along with the instruments for persuasion both by members of the profession and the public.

11. Why code of ethics should exist in banking sector?

Because of the following reasons:
·   Unethical economic practices,
·   External moral and political control over the members,
·   Regulating the business practices and

lobbying for privileges and rights for the profession,

· Adherence to professional ethics as contractual obligation.

12. What is Code of Ethics?

Code of ethics is a set of rules or precepts designed to induce concern that will maintain public confidence. The practicability of the rules of ethics in the banking sector is shown firstly by the manner in which the rules were developed. It should he a code of public relations document in banking sector. Relationship may shift very rapidly in the business world, therefore, alone the legal codes cannot fill the gap for which moral standards are needed.

The moral standards are maintained by God fearing people. Therefore, spiritual values must also be inducted into the code of ethics. To maintain the dynamics of professional life and to institute reforms in institutional framework is to foster the observance of professional ethics.

Codes of business ethics for banking professionals must be related to the following norms:

· charters,
· Contracts,
· covenants and
· consistent with other instruments and
· related ethical priciples discussed in previous sections.

Now the question arises, how to improve performance through integrity, trust and commitment? Ethics of international banking business offers social contract between productive organizations and society. We are to see how business and ethics are to be connected and integrate the teaching and practice of management more closely with ethics and the humanities. A scholarly standard is to be set for banking business ethics. For this purpose, following points may be included in existing code of ethics:

(a) Wise executives worry more about invisible mistakes.

(b) Love the organization that does not love you back.

(c) Be obedient to your commitment.

(d) Keep such management and operational style that doesn't spoil organizational culture.

(e) Do the job with total concentration. Focus yourself to peak performance empowering work force.

(f) Learn from others how to improve business

(g) Give peak importance to vision, values and belief.

(h) Proceed with honest and honorable intention to win everybody.

(i) Be practical as well as generous in your ideals.

(j) Keep your eyes on the stars and keep your feet on the ground.

(k) Remember that hard work, intelligent effort and self-mastery are essentials to success.

(l) Character in the long run is the decisive factor in the life of the individual.

(m) Develop your employees deeply committed with missionary spirit.

(n) Create an environment that stimulates innovation and professional creativity.

(o) Set priorities to focus business efforts on those resources that provide greatest potential return and service excellence to clients and customers.

(p) Adapt yourselfto changing market place.

(q) Time is both a valuable resource and a fixed commodity.

(r) Speed up and attain competitive advantage.

(s) Evaluate the global banking business from moral perspective.

(t) Keep yourself updated by having necessary information and conducting inquiries about client's and customer's financial position.

(u) Safeguard against unintended overdrafts.

(v) Avoid frauds.

(w) Avoid negligence and remember that the banker may be deprived of the statutory protection provided to a Clearing banker.

(x) It is a banker's duty and moral obligation to maintain secrecy of the customer's account. This obligation shall continue after the event of closing the customer's account. The disclosure of secrecy may be under compulsion of law or public duty in the best interest of the bank.

(y) Never deviate from legal procedure.

(z) Quality work, internal communication and training must be integrated into everything you do.

13. Recommendations

1. Extensive programs may be designed and formulated in the field of business ethics, economics and morality for bankers during training at banking institutes. We must also build in-house ethics programs, workshops, seminars and short term training in ethical philosophy related to banking business and guides the employees.

2. All positive steps should be taken to organize the moral education. Try to update legal

and ethical codes by making modifications in the existing codes in view of moral values.

3. As the survival of the fittest is the spirit of banking business, therefore, professional tips, knowledge, skills and business expertise be transferred and the banking management kept fit and smart for modern banking competence using advanced computer software.

4. While dealing with employees, policy makers and customers are qualified, artful, trained and experienced.

5. The rules must to based on and derived from defensible moral philosophy keeping in view the dynamics of business.

6. People of good will and ordinary conscientiousness are to be employed so that they accept the codes and ideas with authenticity.

7. The public representatives must be able to avail the assistance of intelligentsia from outside the particular occupation.

8. If a person has wrongfully been allowed to enter as employee, management must have the right to fire the misfit, if he is not: contributing efficiently towards achievements of pre-settled goals of the organization.

9. The standards of employee's competence, conduct, the educational qualifications and experience should be preferred over political pressure for appointments in the banking sector, especially for technical staff.

10. Many organizations in the banking sector have policies which prohibits employees from accepting gifts; so that the clients are not put in awkward position by violating company policy. A client should use the services because of the value that is provided and nothing else.

11. Measures should be taken to restore traditional values to inject trust and integrity into banking business practices and relationships for growth of profitable relationships with employees, customers, clients, suppliers and alliance partners and to promote service excellence.

12. Measures should be taken to eradicate unethical norms and malpractices for bridge financing. It is a short term financial arrangement given to the industrialists for their projects and recoverable from the sales of products. For this purpose, political bribe, discretion and element of exemptions work. Discretions and illegal exemptions are killer of every system. The unlimited discretionary and exemptionary powers lead to unethical norms. Lacunas in the policies, which keep the policies-unimplemented and ultimately the circle of exemptions and discretion creates the vicious circle of corruption.

13. During the training courses and specially arranged seminars, employees may be instructed about vistas of job success so that they should be well dressed, be punctual, not take frequent leaves, improve ability, be good tempered person with a decent way of communication and expression, have leading, directing, motivating and graceful role in presentations, create self-confidence to be cooperative and with tactful manners without offending the boss.

Acknowledgement:

REFERENCES:

[1] Hadith 335 Narrated by Zaid Bin Talha ,ibid.

[2] Hadith 87, ibdi.

[3] Hadith 87, ibd

[4] All Quran, Al Ahzab :21

[5] Al Quran, Al Ambia: 107

[6] K. Ali, "A Study of Islamic History".

[7] Marmaduke Pickthal, "Quranic Advices", Taj Company, Karachi.

[8] Al Quran: 24:37

[9] Al Quran: 42:38

[10] Al Quran: 28:58

[11] Al Quran: 17:26

[12] Al Quran: 22:41 and 31:17,3:110

[13] Al Quran 7:31

[14] Al Quran 8:58

[15] Al Quran 15:88

[16] Al Quran 2:283

[17] Al Quran 3:161

[18] Abid Ahmad: Shaan e Muhammad (PBUH)

[19] Mahmud Zayid[1980] : English translation of Quran; Lebanon

[20] Gerald Segal [1991]: The World Affairs.

[21] Keith BothmaFollow (2018), " 5 Most Common Types of Banking Fraud and How Digital Transformation Fights T h e m , " Microsoft.:https://www.linkedin.com/pulse/5-

[22] Abdul Fadl, Mirza [1980]: "Saying of Mohammad, peace be upon him": National Hijra Centenary Celebration Committee, Pakistan.

Desktop Based: Off-line Information Retrieval System

Syeda Binish Zahra [1], Syed Muhammad Shabih-ul-Hassan [2]
binishzahra@lgu.edu.pk, binishzahra@ncbae.edu.pk, shabih@pac.edu.pk
[1] Department of Computer Science, Lahore Garrison University
[2] Professional Academy of Commerce

Abstract:

Information retrieval is rapidlydeveloping field and there are many changes are introduced day by day in traditional techniques for IR. IR system is intended to evaluate examine and accumulate the sources of information and get back those that match user's requirement. The need of today's fast moving life is to get maximum information but within minimum time. For getting maximum information in minimum time requires more efforts. The main functionality of IR is to provide access of documents (that document may be in collection of thousand, or millions). With the help of an appropriate description, user can recover any one document from a collection of documents. In this paper I describe my IR system which retrieves information from any directory and this information may be in terms of image, audio or in text form. The selection of good features also allows the space, time and costs of the retrieval process to be reduced. Two documents may be considered similar in this system if they have same name and places in different folders or directories. To explore the retrieval process from that system, I used Apache Lucene with JAVA implemented in IntelliJ.

Keywords: Information Retrieval (IR), human-interaction

I. Introduction:

The task of Information Retrieval is to find appropriate document from a repository. IR system typically accepts a query as a string of words and returns a ranked list of document from that repository [1]. With the help of good human-interaction methods with systems, an information retrieval system can give the ease-of-use of given system even if it having two distinct features. The mean of two distinct features are natural language interpretation and visual interface. Visual interaction with system for local search is done by visual interface and linguistic interaction with system for global search is done by natural language interpretation [2]. It is beneficial to use visual interface for IR systems because by using this user can create expressions of query easily, and can consult with system and even react with the system easily. Most tools are devising for IR to aid information. Many tools are already available to support electric mail, data analysis, manipulation of spreadsheet, preparation of text and even some tools can edit and playback audio\video data [3]. However, for retrieving and sorting above mention information; tools are remaining primitive. In 1951 Bagley's suggested that searching 50 million item records each containing 30 index terms would take approximately 41,700 hours. Since 1960s and 1970s major IBM innovations are introduced such as IBM desktop filing system and free-text search systems. But these systems deals with only information explosion. Our goal is to improve the user ability of retrieving information through their system more easily and also within minimum time slot. We design a system for retrieve the information through directories of user system and that system handle easily the process of finding information of their relevant query. For design such type of IR system we use JAVA in IntelliJ and include libraries of Lucene; the well-known open source engine, for diverse retrieval experiments and in different ways.

II. Proposed System:

Database systems are design to store the entire information in the form of data (all data in documents is structural), but IR system use simpler model than any database system used [5]. IR organized information as a document collection even these documents are not in structure form (no schema).IR locates relevant documents, on the basic of input (enter by user) such as finding documents containing the words "database system". When user wants to retrieve information, then user temporarily enters into a new world. This world is confine from the rest of their computer environmentwhen user returns; they are with result of their requested query. In given system a short query in form of number will do a remarkable job of searching material which is relevant to the query. In this paper we discuss the algorithms that we use in implementation of our IR system. In which we test our system by given it different queries. In final session in this paper our system will generate results of the given system.

a) How our System Works:

Our system offers a full data retrieval mechanism base on their directories numbering. User just enter a desire number of their directory (that numbering is assign to directories by our system) to extract their desire information. After user enters their desire directory number, the system will return all sub-directories of that selected directory. Our system will return a list of all folders which that directory contains.
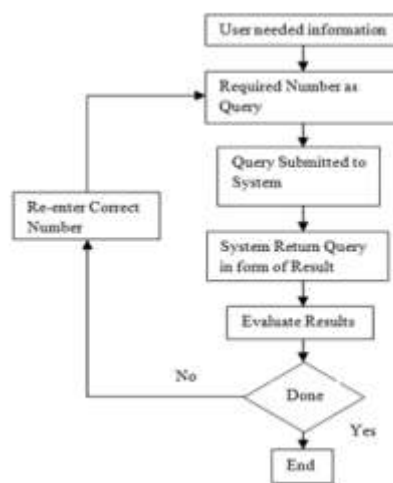

Figure 1: Standard Information Retrieval interaction Model

b) Feature of Our Proposed System:

Our system can easily be supported any new directory that added at run time (a new folder or even a new USB data). For example if we attached a new USB device with our desktop system and that device have multiple folders in it. If we want to retrieve information from that device, our system will also handle those queries too with proper and accurate retrieval of information.
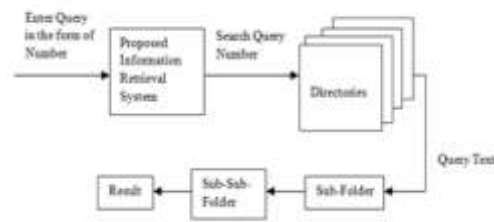
Figure 2: Information Retrieval Proposed Model
It is our system ability to index all types of documents such as plain text documents, audio\video files, Adobe PDF files, Microsoft Excel, Word, PowerPoint files, HTML documents, Web Documents (those web documents that are store somewhere in our system directories) and any image file (jpg, tiff, png and gif format images). Each term extracted from a directory has three fundamental properties
1.   Actual directory
2.   Theposition at which the number assign to subdirectories
3.   The fields in which that query occur.

c) Index Structures [4]:

Our system index consists of 4 main data structures:
Lexicon:
The terms and term id (a unique number for each term) are store in lexicon, along with the global statistics of the term (term document frequency and global term frequency) and the offsets of the postings list in the Inverted Index.
Inverted Index:
Thepostings lists of a term are store in inverted index. In particular, for each term, the inverted index stores:
ü   The document id of the matching document; and the term frequency of the term in that document.

Document Index:
- The Document Index stores
- The number of document (an external of the document's unique identifier),
- The id of document (internal document's unique identifier),
- The document length in terms of tokens;
- The document offset in the Direct Index.

Direct Index:

The Direct Index stores
- The terms and term frequencies of the terms present in each document.
- The main purpose of the Direct Index is to efficiently handle query expansion and also provide easiness.

• III. Implementation And Results:

To retrieve the whole data from specific folder, our system use given algorithm.

```
private ArrayList<File>getAllFiles(String path)
{ArrayList<File> files = new ArrayList<>();
Collections.addAll(files, (new
File(path).listFiles()));
intsize = files.size();
for (inti = 0; i < size; i++) {
if (files.get(i).isDirectory()) {
files.addAll(getAllFiles(files.get(i).getAbsolute
Path())); } } return files;
}
```

The following figure shows the implementations of proposed system and their corresponding results.

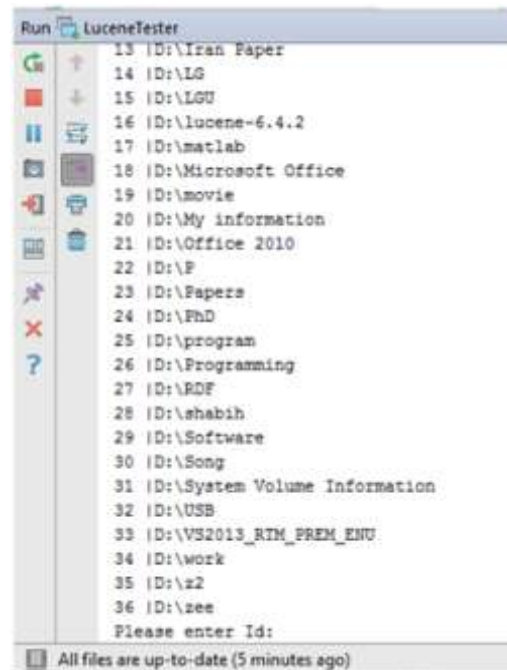Figure3: Results after running the code

Figure 4: Results after enter the id 1, which represent D directory

Figure5: Results after enter the id 24, which represent D directory folder PhD
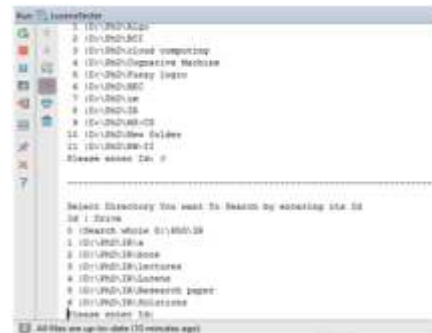
Figure6: Results after enter the id 8, which represent D directory folder PhD, sub-folder IR

Figure7: Results after enter the id 5 which represent D directory folder PhD, sub-folder IR, sub-sub-folder Research Paper
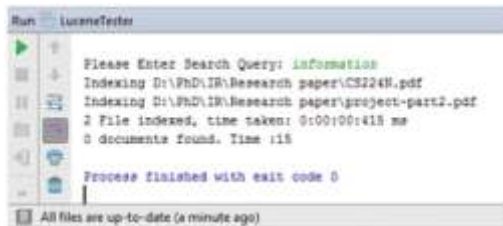
Figure 8: Results after enter the id 0, and then Query "information".

IV. Conclusion:

Our goal with this work is to stimulate new thinking about how to retrieve data through directories in a fast manner. In the future, we hope to evaluate our system with just entering required query and given query will return as output of system with proper data path at only 1 time. With small changes our implementation can be improved and overall retrieval tools will replace by our designed system.

V. References:

1. Hsieh-Chang Tu, Jieh Hsiang, "An Effectiveness Measure for Evaluating Open Retrieval System" in workshop of OSIR (Open Source Information Retrieval)' 06 Seattle, USA

2. Thomas Erickson, Gitta Salomon, "Designing a Desktop Information System: Observations and Issues", in Human Factors in Computing Systems: CHI' 91 Proceeding. ACM: 1991.

3. Michael G. Lamming, William M. Newman, "Activity-based Information Retrieval: Technology in Support of Personal Memory "

4. IadhOunis, Gianni Amati, VassilisPlachouras, Ben He, Craig Macdonald, Christina Lioma, "Terrier: A high Performance and Scalable Information Retrieval Plateform" in workshop of OSIR (Open Source Information Retrieval)' 06 Seattle, USA

5. T. Grust, M. van Keulen, and J. Teubner. Staircase join: Teach a relational DBMS to watch its (axis) steps. In Proceedings of the 29th Conference on Very Large Databases (VLDB),2003.

Causes, Prevention and Law Concerned With Cyber Child Pornography

Syeda Marrium Nizami[1], Gulfraz Naqvi[2], Tayyaba Sultana[3]
Lahore Garrison University
mariyum002@yahoo.com[1], gulfraz.naqvi@gmail.com[2], tayyabaanwar66@gmail.com[3]

Abstract

One of innocent creature of this universe is children and sexually harassing innocent children under the age of eighteen via internet is called child pornography. It is an alarming cybercrime today. People involved in these acts use animated cartoons, films, videos, images and magazines to satisfy sexual desires. This article describes the types of child pornography, associated risk factors and preventive legal measures adopted in Pakistan and worldwide. Access to internet and uncontrolled usage has led to increased child molestation cases. Paedophile cases in Pakistan led to the formulation of new laws under the United Nations Children's Rights Convention (UNCRC). Role of Federal Investigation Agency (FIA) and Non-government Organizations (NGOs) is also discussed. The Source siteswhere the exchange of porn material occur is dark web, its usage and how law enforcement agencies trying to shut it down are made clear here. The tools and techniques helpful in the investigation of child pornography are briefly discussed.

Keywords: cyber crime, investigation, sexual harassment, society, magazines, animations, Children, Pornography.

1 Introduction

Presently there is no definite explanation of child pornography except different reviews of different states. The federal law of US explains that a child under the age of 18 is included in the explanation of child pornography films, animations, magazines and other sexually unambiguous conducts. Sexual intimacy, barbarity, masturbation, and 'lascivious exhibition of the genitals or pubic area' are included in sexually unambiguous conducts. Different explanations are defined by different states so child pornography is difficult to specify in one-liner definition. [1,2]

Masterminds of Child pornography may use child pornography to justify their sexual urges or to blackmail victim, many pornographic criminals are motivated to collect the material for sexual encouragement and for sexual inspiration. So the cases found are not only of child intercourse, exchange of images but many others types that makes for law enforcement agencies difficult to identify child suffering from child pornography. [3]

2 Legal Definition - Child Pornography

Child pornography is basically categorized in two main groups that include different forms of material like images and media. According to the federal law any viewable representation of sexually unambiguous act is considered as child pornography [4]

Visual illustrations containing child pornography are deemed illegal under federal law. These may include photographs, videos, digital downloads, and images produced to depict an actual minor, undeveloped film and video and electronically stored data. Some states consider age of consent to be younger than 18 years old, any depiction of a minor engaging in sexually unambiguous act below 18 year of age is unlawful.[5]

3 Distribution Methods of Child Pornography Material

There are many different acts that are

criminalized when it comes to child pornography and many of these cases allege that the defendant did one or more of the following:

· Downloading images;
· Uploading images;
· Sharing images online;
· Possessing image files on a Smartphone or computer;
· Sending image files via email, text messages, or other online messaging;
· Soliciting minors to participate in making child pornography, often online;
· Arranging meetings with a minor online and traveling to the planned location.

3.1 Risk Factors

Like many other terms, people often make references to child pornography without fully understanding exactly what it may mean in the eyes of the law. Child pornography is increasingrapidly day by day that can destroy generations. It is important to have an understanding of this because there may be things that could be considered as child pornography that they are not aware of.

Department of Justice in USA explains,pornography of children as defined by federal law may include a variety of things. It is also an area that is not protected by the First Amendment. Videos and photographs, even those that are not developed, as well as files stored in electronic format may be pointed to as child pornography. Images or videos of children less than 17 sexually abused or involved in any sexual physical activity is considered as child pornography.

The child does not have to be actually engaged in a sexual act in the image as it is the overall level of sexual suggestion that may lead a particular image to being deemed pornographic. The law makes it illegal to possess, import, distribute or produce such materials. A person may also be charged with a crime for being given such images from someone else. Federal law considers child pornography a form of sexual exploitation and is sometimes referred to as child sexual abuse images.

4 Addiction of Child pornography

Although sexual abusing of children is not a new crime it has been declared as social problems for almost twenty years, current laws enforcement

agencies and social media is giving full attention to the child abusing through Internet. The evolution of sex crimes facility through Internet, along with child pornography control, has given birth to some critical questions regarding uses and abuses of internet and law enforcement agencies are working against these abuses. [6]

The problem of toddler pornography possession becomes idea to be minimized prior to the emergence of the internet. The availability of baby pornography had declined due to felony moves and statutory revisions. But, there is a popular consensus that the internet has made toddler pornography harder. Electronic baby pornography may be received and traded on the sector wide internet, the usage of internet Relay Chat and via other on line sources. On-line, toddler pornography possessors can download infant pornography for collections and distribute photographs to different customers. Internet child pornography ownership cases involve the use of the internet or laptop technology to own and/or gather electronic photos of baby pornography. These investigations gift challenges for regulation enforcement agents within the United States of America and around the world. This take a look at explores law enforcement dilemmas related to defining infant pornography in a pattern of 34 net child pornography ownership cases wherein no perpetrator was arrested. [7] rising in these investigations and gives pointersrelated to those incidents.

The internet has spread out new possibilities for investigation and proof collection in child sex crimes. New era can provide law enforcement retailers with powerful guns and forensic proof frequently missing in traditional child sex crimes. On account that a lot of what takes place on the net leaves a virtual trail, it is viable that this may truly facilitate police investigations of a few toddler sex crimes. [8]

4.1 Issues in Internet Child Pornography

This evaluation tested regulation enforcement dilemmas associated with defining child pornography an internet crime. Regulation enforcement businesses look into those instances using conventional and undercover strategies and might stumble upon precise dilemmas related to the nature of these incidents.

Moreover, there may be significant statutory variant amongst countries and countrywide definitions of baby pornography. There also can be demanding situations related to settle up with truth. These problems can create obstacles for regulation enforcement. [9]

5 Darknet or Darkweb

Darknet means portions of the Internet that are not found by popular search engines. The users remain anonymous; hence the darknet is often a magnet for criminals.
To gain access to the darknet a so-called "Tor Browser" is needed. Because of the high degree of anonymity, the darknet is known as a transit point for drugs, weapons or child pornography. Alternatively the term "deep web" is used. [10]

5.1 Darknet, Darkweb or Deepweb an online parallel universe

The darknet or darkweb offers a menu of illegal activities. It is a platform for the offenders for all illegal activities related to child pornography. Material related to child pornography is not easily accessible through normal browsing but this darknet's anonymity provides refuge to activists. Almost 20 percent of internet is covered by the search engine of Darknet. It is not a simple web but a complex initiative. First of all entering here and becoming part of or member of this Web is not easy. There are proper procedures for that. Then Porn videos and images cannot be easily obtained. Images are hidden in combination of different images; same is the case with videos. Porn videos are hidden between different normal videos. This story is not ending here it continues and these images and videos are in encrypted form and are saved by proper passwords. Encrypting software is required to be developed. Darkweb investigators have been trying to make such software science 2016.

5.2 Investigator's focus: Darknet a platform for child porn

An alarming situation for law enforcement agencies is that this platform is being used by almost 90,000 members all over the world for the exchange of material that shows children being sexually abused. Recent reports and

investigators of Germany and Austria claims that exchange of sexual material has been shut down. The culprits of videos and images showing physical abuse of children are being caught and their systems has been closed. 39 years old German man was caught involved in running illegal site. By the help of this old German's statement investigators are trying to reach the roots of this site. [12]

6 Discussion of few pornography cases

6.1 Child's Physical Abuse images swooped by German Police

In an investigation German police found almost 67 people suspected for their involvement in darkweb. They found systems using IP addresses from where they share sexual material on darknet. The suspects, aged between 18 and 80, were suspected of sharing videos and images of child sexual abuse over the online platform Chatstep. Police said the individuals were identified with the help of "very good cooperation" from the operators of the service. Like other US-based service providers, Chatstep is obliged to tell US authorities about any cases in which child abuse images were exchanged. [13]

6.2 Childs Play for 11 months, Australian police running child porn

A Norwegian newspaper revealed in a major English-language investigative piece that Australian police ran a popular child pornography darknet website for almost a year. [14]
An IT expert investigating for Norwegian tabloid Verdens Gang, known as VG, discovered that the child abuse site called Childs Play was being hosted by police in the state of Queensland. VG made the discovery by exploiting vulnerability in the site's code to reveal its IP address. VG reporters flew to the city of Brisbane to confront police on why they were hosting one of the world's major child abuse image-and-video-sharing sites. They were told details of the operation on condition they did not publish until police were ready. It turned out

police had been running the site for last two years, after the site's Canadian administrator, known as WarHead, was arrested in the US state of Virginia. . [15]

6.2.1 Running multiple sites

Police told VG they had been tracking WarHead, also known as Curiousvendetta, after discovering his true identity when he posted on a technical forum asking for help on code he was using on a separate child abuse site called Giftbox. WarHead reportedly met up with Crazymonk, his co-administrator on the Giftbox site who was also being tracked by police, at a house in Virginia to rape a four-year-old girl on the invitation of a fellow forum user. That was when they were arrested by US police and handed over the keys to Childs Play. Queensland's Argos Taskforce have built up a reputation for infiltrating and taking down child abuse websites and had been investigating the Giftbox site for some time. They also had a legal advantage over most Western police forces in their ability to commit offenses in pursuit of criminals. So a Department in USof HomeLandhanded over the keys to the website to Argos, police told VG. Argos reportedly took over the site and began impersonating WarHead, even posting abuse images until September 2017, when they quietly took the site down. The Norwegian outlet published soon afterwards. During this time police monitored communication on the site and altered code to unmask its users. . [16]

6.2.2 More than one million accounts

Police told the Norwegian journalists that their control of the site and its database of more than one million users yielded information that led to children being rescued and abusers are being locked up. The taskforce was now in the process of sending cases to police around the world. Police told VG that they had a list of between 60 and 90 people on their target list. . [17]

6.3 Australian Doctor Child Pornographer

An Australian Skin Cancer doctor accused of making child pornography. Doctor is behind the bars after accusing that he was involved in possessing and distributing child pornography. Arrested in Canada while he was there as a visiting doctor for some major surgeries.

6.4 Operation Avalanche

Operation Avalanche was held at United States in Fort worth (15th largest city in United States). This operation began in 1999 and ended in 2001. Almost 144 suspects were attested and 100 were found guilty and involved in child pornography. A site Named Landslide Production suddenly became very popular within two years. It seems that this site lists the credit card frauds and information regarding stolen cards. Nothing about Child pornography found unless the user proves identity and encrypt the data found by opening some special sources. There were images and videos of young children being raped and physically abused. The master mind of this site was a husband and wife named Thomas reedy and Janice Reedy. Both of them were self-made programmers of computer science. Their aim was just to earn money and they did. They earned almost $10 million within 2 years with 300,000 subscribers from all over the world. They were caught by the overview of Reedy's hard drive during investigation.

7 Tools and techniques used in the investigation of child pornography

There are some tools used by the investigators in investigating child pornography cases. It involve

·   Visual tools
·   Automated Video Retrieval tool
·   Child Pornography Offender Risk Tool
·   Database
·   Adroit Photo Forensics

These tools work on the principle of image processing. Images found are processed, filtered, enhanced some methods are applied and at the end they are classified. The investigation techniques include identification of distance between eyes nose and mouth for the identification, identification of skin color, identification of body parts and age of the victim and identification of child. Another technique applied is Reverse image search. Computer programmers have developed some software

that detects the image. Reverse image search works on the same principle. When the image is uploaded, the software compares the image with the images used in social websites. And gives the result, if the same image or any image of that child is used somewhere else on any social website. Although all of these are not 100 percent accurate but it is helpful for law enforcement agencies many times.

8 Prevention

8.1 Preventive Legal Measures in Pakistan

The Senate handed a bill in 2016 against the criminals of child pornography, sexual harassment and trafficking. The modification to law generating agencies, that will be enforced on the public after the final recommendation of the president, the age of criminal responsibility also increases from seven to 10 years of age. Under the revised rules, sexual assaults will now be punishable [18]

Similarly, infant pornography, which turned into punishable crime now was previously not mentioned within the law, may be punishable by means of seven years in prison and a satisfactory amount of Rs1.7 million. Pakistani investigation agencies and law enforcement agencies became active, when revealed the exchange and transfer of pornographic material including images and videos from a small village in Kasur district named Hussain Khanwala in Kasur district were created and distributed all over the world. [19] The united states become active in August 2016 by means of a primary paedophilia scandal (Catholic Church Case) when girls and boys under the age of 12 were sexually harassed by the priests of the church.

8.2 The Federal Investigation Agency (FIA) and child pornography

FIA is an investigating agency in Pakistan that took great actionsto curb child abuse. [20]FIA caught a gang in Jhungwith the help of Interpol Canada, who were indulged in the sharing of pornography material. Offenders were arrested with the 60 Giga bytes of data. [21]The data include porn videos images animated films, ads etc. The data found in laptops, hard drives and other storage devices and a professional computer engineering found involved.[14] IP addresses are the main sources that were found helpful in tracing the culprits. Acting upon such crimes improves the livings standards as well as the security of children. [22]

8.3 Punishment for child pornography in Pakistan

There was no law of punishing the criminal of child pornography before the famous Kasur case that happened in 2015.[23] This was the first case of child pornography that came in social media and provided awareness to the people.[24] After this case a law was passed by the national assembly of Pakistan which stated that Child pornography offences in Pakistan can be prisoned ranging from 2 to 7 years including a small amount of money. [25]This punishment was not enough as this crime is increasing day by day so on 11 august 2016 a new cyber law was passes.[26] According to that law Child pornography offences in Pakistan can be prisoned ranging 14 to 20 years including some amount of money. [27]

8.4 Role of Non-Government Organizations (NGOs)

According to a survey law enforcement agencies in UK are paying special attention on the internet related crimes including child pornography. [28]They have arrested almost 1,713 criminals involved in the activities like possession of child, pornography depicting graphic sexual images, possessed images depicting prepubescent children etc. [29]

9 Big issue

Technology has made child pornography more pervasive. The age of computers and social networks and social media has led to a sky rocketing number of crimes involving child pornography.[30] The types of persons who commit this offense often do not have any criminal history that involves the sexual abuse or exploitation of children, nor do they have any prior felony.[31] Access to child pornography can be accessed from the privacy of home computer over the internet. Gone are the days when producing and distributing these kinds of

images was difficult and expensive. Anonymous distribution was not possible.[32] The Explosion in the volume of child pornography is the result of technology abuses. [33]The ease that technology has provided to the people of getting the wanted material on a single click, by millions of people at the same time is the basic source. Apart from this getting such images digital movies, magazines and other material on child pornography is not at all difficult in this social world and that is indeed an alarming situation [34]

10 Working

Many NGOs are working against this child pornography.NGO is a non government organization that work for the welfare of human rights. [35] According to a survey the percentage of this abusing is unfortunately increasing 4.3 percent every year, and that is an alarming situation. The second alarming and shocking thing to be noticed is that according to a report the rate of child pornography is almost 150 in a year aged 6 to 10 comparatively to the rate of child pornography in boys is higher 178 cases in a year.[36] In such situation it is very important to take some serious steps to realize obligationsof Pakistan under United Nation's Convention on the Rights of the Child (UNCRC).[37] Pakistan ratified here in 1990 under the obligation to obey the rules and implement its outline by the cooperation of police, special law enforcement agencies. [38] Establishment of a national commission on child rights, to monitor and coordinate the proper working of implemented, legislation, special departments for the report of progress to the committee.[38] This will provide a jump board to the Children Rights in Pakistan. Every country that is connected in this committee is answerable of the progress after the period of five year. [40]

Many state and federal law enforcement agencies may engage in undercover sting operations in which they pose online as either minors or other individuals seeking child pornography.[41] These operations can result in the search of home, including all smart phones, laptops, tablets, desktops computers, and hard drives.[42] If evidence of child pornography is found, criminal can be arrested and charged with a third or second-degree felony, or even federal

charges, depending on the circumstances and the acts alleged.[43]

10.1 Contact Police and crime investigation agency

The production, possession or allocation of child pornography is a humorless crime. If found any evidence that someone is committing an act that is connected to child pornography, it is best to contact the police first and the second important act that should he done is to inform the nearest cyber crime investigation agency. In Pakistan National Response Center for Cyber Crime Federal Investigation Agency Pakistan (NR3C) is working under FIA specified for the electronic or technical crimes. Every citizen in Pakistan can easily contact to NR3C through E-mail helpdesk@nr3c.gov.pkfor help. Official web site of NR3C can be visitedhttp://www.nr3c.gov.pk/contact.htmlwhere address, fax number and contact number is given for every city. Their aims and objectives are completely described on this link and online complain system is also available for the ease of victim. [44]

11 Conclusion

Cybercrimes such as child pornography is increasing at an alarming rate due to advanced technology tools. Not only Pakistan but my countries are also affected by the hilarious crime and some legal measures have been implemented by law makers to curb this criminal act. The role of private organizations like NGOs and public agencies is warranted to enhance social awareness and to enforce penalties. Better tools and techniques are being developed by the researchers to reduce the rise of this crime but government and public support and cooperation is necessary. Victims should raise their voices without any fear of privacy issues and of being mistreated. Trust should be developed between public and law enforcement agencies.

12 References

[1] https://www.joefriedberglaw.com /blog/2017/03/how-is-child-pornography-defined.shtml

[2] (PDF) Defining Child Pornography: Law

Enforcement Dilemmas in Investigations of Internet Child Pornography Possession 1. Available from:
https://www.researchgate.net/publication/228814315_Defining_Child_Pornography_Law_Enforcement_Dilemmas_in_Investigations_of_Internet_Child_Pornography_Possession_1 [accessed Sep 20 2018].

[3] http://law.jrank.org/pages/11986/Cyber-Crime-Online-child-pornography.html

[4] Biegel, S. (2001). Beyond our control? Confronting the limits of our legal system in the age of cyberspace. Cambridge, MA: MIT Press.

[5] Brown, D. (2002). Pornography after the fall of the CPPA: Strategies for prosecutors. American Prosecutors Research Institute Update, 15(4), 1-2.

[6] Caruso, D. B. (2003, April 6). New catalog will aid in child porn fight. Boston Sunday Globe, p. A11.

[7] Copine Project. (2003). UCC project tracks down paedophiles. Retrieved November 8, 2004, from http://www.copine.ie/press.php#16jan2003

[8] Douglas, G. (2002, April). Jim McLaughlin's secret war. Yankee, 74-82.

[9] Finkelhor, D. (1984). Child sexual abuse: New treatment and research. New York: Free Press.

[10] https://www.dw.com/en/darknet/t-19184846

[11] https://www.dw.com/en/darknet-parallel-online-universes/a-19413906

[12] https://www.dw.com/en/investigators-bust-major-darknet-child-porn-platform/a-39570195

[13] https://www.dw.com/en/german-police-swoop-over-child-sex-abuse-images/a-39764972

[14] https://www.dw.com/en/norwegian-newspaper-reveals-australian-police-ran-child-porn-site-childs-play-for-11-months/a-40865610

[15] https://www.dw.com/en/europol-fbi-arrest-nearly-900-in-crackdown-on-global-pedophile-ring/a-38729883

[16] https://www.dw.com/en/police-crack-down-on-norways-largest-pedophile-abuse-network/a-36458980

[17] https://www.dw.com/en/alphabay-and-hansa-darknet-markets-shut-down-after-international-police-operation/a-39776885

[18] Pornography: Research advances and policy considerations (pp. 235-255). Hillsdale, NJ: Lawrence Erlbaum Associates. Downloaded By: [Thirkers, Queen] At: 12:43 25 July 2007 282 M. Wells et al.

[19] Hames, M. (1994). A police view of pornographic links. In P. C. Bibby (Ed.), Organized abuse: The current debate (pp. 197-203). Brookfield, VT: Ashgate.

[20] Hardy, R., &Kreston, S. S. (2002). 'Computers are like filing cabinets …' Using analogy to explain computer forensics. National Center for Prosecution of Child Abuse Update, 15(9), 1-2.

[21] Holmgren, B. (2002, Spring). Translating science into law: Lessons from doctors, judges, and lawyers about the use of medical evidence in the courtroom: The expert witness. New England Law Review, 36, 593.

[22] International Centre for Missing & Exploited Children. (2002). Campaign launched to attack global problem of child pornography. Retrieved November 8, 2004, from http://www.icmec.org/missingkids/servlet/NewsEventServlet?LanguageCountr

y=en_X1&PageId=1295 International cooperation to tackle child pornography. (2001). The Lancet, 357, 569.

[23] Jenkins, P. (2001). Beyond tolerance: Child pornography on the Internet. New York: New York University Press.

[24] Klain, E. J., Davies, H. J., & Hicks, M. A. (2001). Child pornography: The criminal-justice-system

[25] response. Alexandria, VA: American Bar Association Center on Children and the Law for the National Center for Missing & Exploited Children.

[26] Lanning, K. V. (1992). Child molesters: A behavioral analysis (#98-MC-CX-K002). Washington, DC: National Center for Missing & Exploited Children.

[27] Lanning, K. V., & Burgess, A. W. (1989). Child pornography and sex rings. In D. Z. J. Bryant (Ed.),

[28] Pornography: Research advances and policy considerations (pp. 235-255). Hillsdale, NJ: Lawrence Erlbaum Associates. Downloaded By: [Thirkers, Queen] At: 12:43 25 July 2007 282 M. Wells et al.

[29] Lemmey, D. E., & Tice, P. P. (2000). Two tragic forms of child sexual abuse: Are they often overlooked? Journal of Child Sexual Abuse, 9(2), 87-106.

[30] McCullagh, D. (2003). Supreme Court weighs Net porn law. CNET News. Retrieved October 14, 2003, from http://news.com.com National Center for Prosecution of Child Abuse. (1999). Child abuse and neglect state statues elements: Crimes Number 30: Child pornography. Alexandria, VA.

[31] National Directory of Law Enforcement Administrators. (2001). Stevens Point, WI: National Public Safety Information Bureau.

[32] Norland, R., & Bartholet, J. (2001, March 19). The Web's dark secret. Newsweek, 44-51.

[33] Persson, D. I. A. (2001, December 7). International perspectives on Internet crimes against children.

[34] Paper presented at the National Internet Crimes Against Children Training Conference, New Orleans, LA.

[35] Rosenbloom, A. L., & Tanner, J. M. (1998). Letter to the editor. Pediatrics, 102(6), 1494.

[36] SEARCH. (2001). Combating Internet crimes against children: Recommendations of the Online Child Exploitation and Victimization Focus Group. Sacramento, CA: Search.

[37] Taylor, M., Quayle, E., & Holland, G. (2001). Child pornography, the Internet and offending. The Canadian Journal of Policy Research, 2(2), 94-100.

[38] Taylor, S. (2001, March 19). Is it sexual exploitation if victims are 'virtual'? Newsweek, 44-51.

[39] Tyler, R. P. T., & Stone, L. E. (1983). Child pornography: Perpetuating the sexual victimization of children. Child Abuse & Neglect, 9(3), 313-318.

[40] US Department of Justice. (2000). Use of computers in the sexual exploitation of children (NCJ170021). Washington, DC: US Department of Justice, Office of Justice Programs.

[41] Wells, M., Finkelhor, D., Wolak, J., & Mitchell, K. J. (2003). Internet sex crimes against minors: Law enforcement dilemmas. Washington, DC: US Department of Justice.

[42] Wolak, J., Mitchell, K. J., & Finkelhor, D. (2003). Internet sex crimes against minors: The response of law enforcement. Washington, DC: National Center for Missing & Exploited Children.

[43]   Wolak, J., Mitchell, K. J., &Finkelhor, D.
         (2004). National Juvenile Online
         Victimization Study (N-JOV):
         Methodology report. Crimes Against
         Children Research Center. Retrieved
         March 19, 2007, from
         http://www.unh.edu/ccrc/national_juven
         ile_online_victimization_publications .h
         tml

[44]   http://www.nr3c.gov.pk/contact.html
         helpdesk@nr3c.gov.pk

Importance of Validating Evidence Integrity in solving Crime

Fatima Gul
Lahore Garrison University
fatima.dfrsc@lgu.edu.pk

Abstract

At present time, the crime rate is growing fast and this alarming rise can be controlled by using evidence as a key component. For solving any crime, evidence should be admissible in court of law by protecting its integrity. The weight of an evidence determining the truth of an issue is entirely subjected to the examination and verification of present forms of legal arguments. . It is duty of forensic practitioner to keep the integrity of evidence by not distorting any relevant information, ignoring contrary evidence and misinterpreting any available evidence. The forensic expert should advocate his outcomes with proper vigor and forcefully once he reaches to final conclusion. His credentials must be represented accurately without any misrepresentation and by maintaining his practice and specialty competence. The reporting information should be relevant to the evidence and to then case. Our society and individual can have life changing consequences from the forensic analysis interpretations and conclusions therefore scientific analytical data should be generated via reliable practices based on valid scientific principles and methodology. It is vital to validate any evidence before evaluating its weight. However, the forensic science development is dynamic process from data generation to result interpretation that is evaluated and communicated in such way that can help the court of law to make sound decisions relevant to forensic evidence admissibility.

Keywords: Evidence, Investigation, Forensic, Validation.

1. Introduction:

With the growing rate of high-tech world, the rate of the crime is increasing. The expending law is countering fast growing crime rate persistently. Evidence is collected on the committing of the crime and according to Locard's exchange principle every contact leaves a trace which may be used to show the association between the crime scene, victim and suspects. After examining andanalysing the evidence from a crime scene, the criminal or perpetratorwho isactually responsible for keeping the crime ratio high, are identified. For this purpose present investigation process needs to be improved to support the case proceeding. Various steps are involved in the criminal proceeding cases like those of police investigation with FIR, law enforcement process and the final verdicts. Evidence collected from the crime scene, scientific evidence which is gathered by scientific methods, is of significant value in all stages while providing information and hints in sexual assault cases to the police department, prosecutors and to the court room. Investigators can have many type of evidence to solve the crimes such as trace evidence questioned document, impression evidence, finger print, controlled substances, firearms, toxicology, biological, electronic, digital, testimonial and hearsay evidence. The analysis of forensic evidence can be carried out by emphasizing on team work, using developed forensic skills/tools such as video image, data analyser, cell phone tracker and GPS positioning. However, a crime scene is processed by proper identification, collection and preservation of the relevant physical evidence (4, 5, 6, and 7).

To prosecute criminals in the court room, sound evidence is required. Evidences with forensic value can be analysedto solve crimes by validating its integrity.Validating of evidence integrity is protecting of potential evidence from being destroyed/damaged and to prevent introducing of false evidence in the question area. The evidence integrity can be protected by

maintaining chain of custody and making it admissible in the court of law. Where the role of forensic scientists is vital in criminal investigations and prosecution, they are also contributing to civil limitations, manmade terrorism, natural disaster, and global crime investigation (1, 4, 6 and 8).The evidence becomes inadmissible if its integrity is not proved in the court of law and it becomes questioned even if there is a doubt of being tempered, mishandled and giving it to an unauthorized person's custody. The integrity of evidence from collection till prosecution must be kept safe bymaking chain of custody tamperproof (2, 3, and 6).

Best practice of forensic science can be achieved by additional scientific examination and validation. For this purpose, scientific methods/practice must be reviewed and updated periodically to the recent development in the relevant field (15). Validity of the evidence integrity can differ depending on its nature, forms, types of crime scene and the scientific methods used to process relevant crime scene. Generally relevant forensic evidence is perceived to be of a weapon or a bullet found in the body of victim or the pattern of blood in a homicide case. However, to solve the crime, any particular piece of object can be proved crucial as part of physical evidence. Therefore, any type of evidence found at the crime scene can be considered as a physical evidence to help in investigating the crime (2, 3, and 8). Here in this study, a general over view of importance of validating evidence integrity will be discussed.

2. Review

Science and law have different nature, science seeks to develop our knowledge about natural world and the law ensures the public safety and the justice to be served properly. For achieving these goals, with the passage of time science became available as tool to the legal system. The result and conclusion obtained from forensic data can have life-changing consequences on our society. So therefore, it is important to use reliable scientific methods built from valid scientific principle and methodology to process any analytical data (16, 10). In some countries, the judge is responsible for assuring the admissibility of scientific evidence, as the advancementin forensic analysis is

dynamic, from data collection to its interpretation. As a result of this, the weaknesses in the forensic methodology due to advancement can arise issues for the legal practice. To have sound decisions for the admissibility of forensic evidence in the court of law, the advancement in the scientific methods and technology must be evaluated and validated (15).

Validation can be performed to assure the efficacy of forensic evidence and methods used. The term validity defines theextent at which a scientific test is intended to be measured or more specifically the extent at which the tests are best justified made by the examiner. Therefore validation can be defined as a process of evaluating the efficacy and reliability of the procedure carrying out forensic analysis. Reliability which is the measurement consistency can also be considered as an aspect of validity. Thus any type of measurement inconsistency indicates threat to the validity of the methods which should not be confused with internal validation and that is checking performance of test methods developed. The location and environmental conditions can also compromise adversely the efficacy or quality of the test measurements and results. The internal validation is different from developmental validation as the internal validation is carried out by forensic science service provider and the developmental validation is responsible for acquisition of the analytical data and assuring of the sensitivity, accuracy, stability, reproducibility and determination of the forensic test method conditions, limitations and evaluation. Internal validation is the accumulation of analytical data that is used to demonstrate the performance of methods and procedures established in the laboratory (13, 10, and 11).

2.1 Validation Process:

The process of evidence validation lies between the development of the test methods and strategy, its scientific acceptance and evaluation of independent reliability and specific purpose relevancy. The evidence validation primary focuses on the performance of alternative test methods developed, their detailed study and processing the analysis of data resulted. Test method validation criteria were firstly originated and developed by the European Centre for the

Validation of Alternative Methods (ECVAM) and European Chemical Bureau (ECB). These criteria were subsequently applied for the development and validation of the alternative test methods and procedure of US Interagency coordinating Committee and Organization for Economic Cooperation and Development (OECD). Now test methods validation according to ECVAM, EC or OCED criteria and principle is widely accepted as prerequisite for test methods and strategies. These criteria and principle for test method validation are periodically updated and also available for peer review (13, 9, and 12).

The validity process constructed must be based on evidence. And test method developed must undergo evidence based validation process for assuring the interpretations obtained from the results to be valid. As validity process is multilayer concept, so different types of the evidences are required to support validity test results and relevant claims. The process of validity is done by using both quantitative and qualitative research data and methods. All the evidences in the validation process should be collected, analyzed and reported methodically. The validation process is of two events i.e. before the test and after the test event. Before the test event includes the design and development phase while after test event is of obtaining data from trialing and live testing phase. As said earlier that validation is a dynamic process that may have different test methods so therefore a list of step by step requirements/guide relevant to evidence validation can be provided as following(13, 15 and 10);

2.2 Evidence collection:

Clearly, the selection and collection of different types of evidence, their source of identification either they are containing required material or not and of what quality and are they relevant or not,are of crucial issues. Once the collection of evidence is complete, it is questioned that the evidence collected is according of the perquisite criteria without any biasand the evidence is relevant to the test methodology or not. Beside these concerns, it is also important to know that how the data was interpreted, the equipment and materials used in the test methodology are relevant or not. As it is obvious that the crime scene remains intact, the crime scene

investigator should evaluate consistencies and inconsistencies during the initial assessment of the crime scene, may be helpful to the continuation of the investigation. There are special cases where the victim is neglected, so additional crime scene need to be considered and for this purpose, the evidence must be collected and photographed using strict and approved methods in order to maintain the evidence integrity by avoiding contamination. The chain of custody should be strictly enforced during evidence collection.

Group of experts, such as information technologist and scientist, should control the collection of evidence as they are familiar withmethodology.Unless the data/evidence has been analyzed by subjecting to the formal set of criteria, eitherrelevant or against the methodology, should be initially acceptable for the review. The criteria of inclusion and exclusion should be cleared before the data retrieval, transformation and analysis in report made by validation assessment. The information produced as a result of test methodology and the reference data for the assessment of test performance should include following points;

· Data should be relevant to the evidence of interest e.g. biological or other targeted object.

· Detail description of the reference material including its source and quality for the assessment of proposed test methodology.

· There should be access to the all-raw and transformed data, and to the original laboratory record.

· Data quality assessment should be performed to check if they were resulted according to good laboratory, clinical or cell culture principle and practice, and also to ensure the internal and external quality control test.

· Provide the reason of not using any relevant data.

The evidence collected should be provided in the form of peer reviewed publications. However, in some circumstances, the company report can be acceptable if they provide it as public domain after the conclusion of the study. While releasing the data, care should be taken to make bias free publication with positive findings. Any useful information should be taken into account including human response with its relevant

nature and extent.

2.3 Weighing of Evidence:

The weighing of Evidence (WoE) can be described as the consideration of the situation which is used to find out the certainty or uncertainty of the evidence collected, supporting the one side argument is greater than the other side or not. Every individual make weighing of evidence process which are used in different circumstance including educational, commercial, health, scientific and of legal interest. WoE is used in scientific literature and review publishing and is of following characteristics;

· WoE is metaphorical which refers to collection of an unspecified methodological interest.
· Methodological approach can be used to establish interpretative methodology which indicates that the WoE describe the methods that use quantitative weights of evidence or it examines all the subset of the evidence.
· WoE evidence is also used for the conceptual frame work.

The term WoE is of great value as it indicates towards different questions such as, whatmethods of interpretation areused and how they were applied to scientific evidence? These kind of questions are important in finding out the validation procedure which can be used for evaluation/establishment of scientific method's validity for a particular purpose.

2.4 Assessment of Validation:

Validation assessment can be made by providing clearly statedoutcome, supported by detailed and reasoned arguments which should be available to the public. These outcomes or conclusions are mainly of three types that can be used for resolving the uncertainty of the methods developed and validated for a particular purpose as follows;

· The evidence is consistent and sufficient with relevant test methods and its stated purpose for which it should be accepted.
· The evidence is inconsistent and insufficient with relevant test methods and its stated purpose for which it should be accepted,

and the additional evidence of quality and quantity type should be assessed further.
· The evidence is sufficient with relevant test method but not reliable, and should not accepted for the purpose stated.

The result of the assessment should be available in peer review, also should be submitted to the validation assessment sponsors and to the relevant personals for the transparent and independent assessment as a whole peer review (9 and 12). All the stakeholder should respect the evaluation of the evidence validity with underlying forensic methods that can be used for legal and scientific purpose. The Department of Justice and National Institute of Standards and Technology (NIST) submitted two outline in support of National Commission on Forensic Science as follows;

· Research supporting the development and dissemination of methods, technical guide lines and standards should be conducted the measurement of forensic science.
· The existing forensic science practices and standards should be selected as appropriate test and validation.

NIST is distinguished as one of the internationally recognized trusted scientific and technical laboratory. The reviews submitted by the NIST can be of resolving the gap between the scientific validity and admissibility of the evidences (10).

3. What if Evidence is not validated?

Where the forensic science has been advanced and developed, problems in its some aspects also have come to light during the recent years such as false conviction by using faulty firearms and bite mark analysis, incorrect identification of fingerprint and forensic lab misconduct. By knowing these shortcomings, it can be more effective to put forensic science forward by recognizing which parts are scientifically valid or not and to have more research to validate it. Scientific validity should not be confused with admissibility but yet this difference is not clearly understood by those who are involved with court proceedings. Such as bite marks validity as identifier is still admitted by some courts but now it is reversed as DNA analysis is helping in false conviction. Therefore, the concept of admissibility falls back on past decision as the

lack of scientific validity in forensic investigation has been recognized (15).

Similarly, improper handling during a forensic investigation can result in unsuitable analysis in solving a crime and lead to false conviction. A high quality database management can be achieved by maintaining some important conditions such as correct sample collection, chain of custody, safe transportation and storage of evidence, and proper standard analytical procedure. Following these standards leads to high level of confidence which results in good credibility and reliability. Credibility is often compromised due to contamination problems as strong and severe reason. Contamination chances are high at any level of analytical procedure as in sample collection, chain of custody, transportationand analysis. Contamination can be minimized by estimating errors but never can be prevented (1, 14). Analytical procedure of forensic evidence that how it was analyzed is now questioned in many highly technical testimony trails.Questions like these can lead to case failure technically if not prevented and eliminated. Evidence collected from the crime scene such as shoeprint, latent finger prints and biological samples should be identified and explained before they go to the court proceedings else they can lead to false conviction as in the following cases;

·	Adam Scott twenty year old was charged in a rape case of a woman as his saliva was reused which the police collected from in a previous case from plastic tray during a street fight. This evidence should have been disposed of even the case was ended. This faulty case proceedings by a worker jailed Adam Scott for five months before the mistake was recognized.

·	A woman who was murdered in 1997 was taken to the laboratory for forensic analysis. The investigator found another women profile as suspect to be her killer after searching under her nails. The woman was thought to be suspect or killer, was also murdered herself three weeks prior to the incident. The investigators were confused due to not finding any correlation between these two murdered women but finally they came to conclusion that during taking DNA samples, they used same pair of scissors for cutting their nails though they were washed but lead to contamination due to presence of DNA.

·	Similarly, Cory Carroll was charged for murder in a head on collision in which one person was killed, he and two others were injured. In this case the evidence was determined inadmissible and he was released. His urine sample was found with high levels of THC (tetrahydrocannabinol) which were considered invalid as these type of results were never sent to the state lab for further confirmation. The case was dismissed as there was not enough evidence against Cory Carroll.

The evidence integrity also has been compromised due to scientific misconduct which involved the fabrication of data and falsification of the results as in following cases (14);

·	Larry Benedict, 45 year Xerox engineer was sentenced for four years in trafficking of child pornography case by the federal judge. The evidence in this case was electronic which found to be helpful towards his innocence after he hired a computer expert who proved it that the evidence presented in court was allegedly tempered before or after in government custody.

·	Jodi Arias in Arizona was sentenced to death after she found to be guilty in Travis Alexander murder case. Similarly, she also hired a computer expert to examine the victim computer as thousands of files were deleted during in custody of Mesa police department (6). Development in the scientific experimentation and observation contribute to the advancement in the field of research with peer review and publication, and theories. Science is always attracted towards values and forward progress. In field of science, what is known to day may be proved invalid tomorrow that would be discussed as a progress and innovation. This may be of current controversy and concern that the forensic discipline that were not known in the field of science may go through crucial scientific methodology and review check list but it is not considered to be of no scientific value or invalid. Despite this, they are to be demonstrated as such and for this purpose, they must pass the scientific enquiry of present time. For example, since 1900s the finger prints have been used in legal processes as source of identification. In 1911, the decision was made to admit finger prints as evidence in the court by cooperative and judicial system that did stem from 21st century scientific standards. The scientific analysis that applies to every forensic science methods from DNA evidence to pattern evidence, is an ongoing process that is not fixed once done.

4. Conclusion

It is concluded that the validation process for every scientific discipline is essentialincluding forensic science. The importance of validation should be described by means of parameters in the methods development. Validation should include the context and purpose of why it is being validated. For validating process, case by case approach is followed as different kind of evidence have different level of value in overall assessment of validation. The evolution of the evidence validation involves the probability, relevancy, stability, capacity and strength of evidence. Validation assessment can be made by use of available information including systematic, independent and transparent review which can conclude that test methods/practice are reliable foe intended purpose without a dedicated practical study. After assessment, the outcome should be clearly stated by means of support and reasoned detailed argument available publically. The assessment outcome should be published in peer review journal by submitting it to the validation assessment sponsor as whole study i.e. from methods design, evidence collection, and chain of custody, validation, assessment and final reporting.

5. References

[1]   T.Magalhães, R.J.D. Oliveira, B. Silva, F. Corte-Real and D. N. Vieira, "Biological Evidence Management for DNA Analysis in Cases of Sexual Assault," J. Scientific World., pp. 1-11, 2015.

[2]   Chapter (5) Forensic Analysis, "Challenges and Choices for Crime-Fighting Technology."

[3]   H.C. Lee and E.M. Pagliaro, "Forensic Evidence and Crime Scene Investigation," J. Forensic Investigation., vol. 1, pp. 1-5, 2013.

[4]   J. Ashcroft, D.J. Daniels and S.V. Hart, "Forensic examination of digital evidence," US department of Justice, pp. 1-88.

[5]   Chapter (1) over view of Evidence, Jones and Bartlett Publisher, pp. 1-48.

[6]   M.S.M.B. S. Saleem and R. Zulqarnain, "Protecting digital evidence integrity and preserving chain of custody," J. Digital forensic, security and law., vol. 12, pp. 6-30, 2017.

[7]   Patel, Dr. S.L. Vaya, J. Nirali, V. Margi and S. Sharma, "Review of the role of forensic evidence in criminal justice system," J. Multidisciplinary Studies., vol. 3, pp. 2-246, 2016.

[8]   T. McEwen, "The Role and Impact of Forensic Evidence in the Criminal Justice System," U.S department of Justice, 2010.

[9]   M. Balls, P. Amcoff and S. Bremer, "The Principles of Weight of Evidence Validation of Test Methods and Testing Strategies," National institute of health., pp. 603-620, 2006.

[10]   "National commission on Forensic science," U.S department of Justice, 2016.

[11]   "Draft Guidance: Digital Forensics Method Validation," 2016.

[12]   B. Budowle, S.E. Schutzer and A. Stephen, "Criteria for Validation of Methods in Microbial Forensics," J. American Society for Microbiology., vol. 74, pp. 5599-5607, 2008.

[13]   S. Saleem, "Protecting the Integrity of Digital Evidence and Basic Human Rights during the Process of Digital Forensics."

[14]   B.E. Turvey, "Evaluating law enforcement and forensic science cultures in the context of examiner misconduct," Criminology Department Faculty of Humanities and Social Sciences

[15]   Bond University, 2012.

[16]   S. Bell, "Now who will push ahead on validating forensic science disciplines?," Phys.Org, 2018

Spammer Detection: A Study of Spam Filter Commentson YouTube Videos

Rafaqat Alam Khan
Lahore Garrison University
rafaqatalam@lgu.edu.pk

Abstract:

This paper presents a methodology to find out the spam comments on YouTube videos. The purpose of this research is to find out the comments of those spam users, who comment for their own promotional intentions or to detect users whose comments that have no relevancy with the given video.The monetization policy introduced by YouTube for its user's channel and advertisement of different ads on YouTube videos has attracted a large number of users. This increase in a large number of users has also lead to an increase in malicious users whose job is to create automated bots for commenting and subscription to different YouTube channels. These malicious users'comments hurt the channel publicity and also affect the normal user's experience. YouTube is also working on this issue by using different methods to limit these kinds of automated bots malicious comments by blocking those comments. These kinds of methods are ineffective so far as spammers have found out different methods to bypass those heuristic approaches. Different machine learning approaches provide somehow better classification accuracy with the introduction of new approaches to solve it better than that. In this work, different techniques used for classification of spam comments with those of normal user comments to improve the classification and recent trend going on in this area are briefly analyzed to tackle this major issue.

Keywords: Spam, YouTube, Classification, Comment, Social Media

1 Introduction:

YouTube a video sharing website was started back since 2005. In 2006, Google bought YouTube and nowadays it is on Google Subsidiaries. After it came under the umbrella of Google YouTube growth as a YouTube, video sharing has increased significantly. The users using YouTube can create their own Gmail login and through this Gmail login, they can create their own channel. Once the channel is created, YouTube allows users to publish his own video, rating, Comments, likes or dislikes, reporting and subscribing to your favorite channels. According to the recent statistics of YouTube, it has achieved the marked of 1 billion users log in. The global research [1] statistics of YouTube claims that around 1.9 billion users visit within a month, watching billions of hours per day and in turn generating billion of views per day.The 70 % of these watch hours come from mobile devices. Around $ 2 billion dollars, YouTube has paid over the last five years to its YouTube channel owners.

Out of different YouTube features for its channel users, YouTube commenting feature is one of the important features in which users can able to comment on other channels uploaded videos. This powerful feature of YouTube allows the interaction of YouTube channel owner with other users. Introducing of such feature has also allowed other malicious users to promote their promotional content to know as a spam comment. These spam comments are usually irrelevant and are generated by mostly automated bots. The capability of bots to perform spam comments on YouTube videos has been discussed in [2].
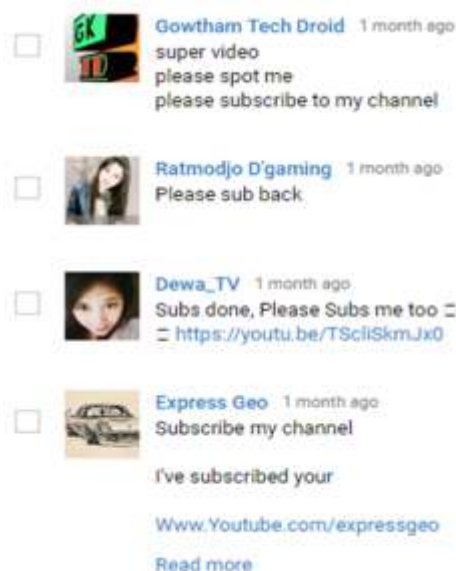
Fig.1. Example of Spam Comments on YouTube Videos A recent report [3] regarding spam comments has been published in BBC, in which YouTube claims that during the period of July-September 2018 they have almost deleted 224 million malicious comments out of billion of comments published in this 3 months period. They have also accepted that this a really hard work to tackle this issue and they are investing in this area to get rid of that malicious content.

The spamming can be basically divided into two categories i.e. link-based spam and channel-based promotion spamming. Link-based spamming is that type of spamming in which users comments different links comment on published YouTube videos, on which other users click than it redirected to malicious webpage's. While on the other hand, channel-based spamming is used for promotion of user channels to get subscribers, views, and likes for his channel as well as for his channel videos. YouTube generally use comment blocking [4] on the basis of HTTP URL, but this form of filtering has to lead the spammer a new approach to publish their links by insertion of white spaces in their link and instead of .com they typed dot com in the form of text to publish their link on the given video. These problems have to lead to the machine learning approaches to resolve such issues. The contribution in this research is finding out of better spam classification using different machine learning approaches with high F1 score and better classification accuracy.

2 Related Work

Chowdhury Rashid et al. [5] generated a lift chart by using three different data mining models. This lift chart finds out the lift score when compared to a random guess. The predicted probability for the three different data mining models i.e. Naïve Bayes, Decision Tree and Clustering is calculated. From the result, they have concluded that for most cases Naïve Bayes and decision tree performed better than that of clustering.

Tulio C. Alberto et al. [6] used different classification algorithms i.e. Naïve Bayes, Decision tree, SVM, Random forest and logistic regression on five different datasets. They have achieved a confidence level of almost 99 % on all these classifiers. Based on these classification results they have developed their own online tool known as TubeSpam that automatically detects the spam message on the fly.

SaumyaGoyal et al. [7] spam message detection on real twitter social media dataset is applied using KNN and decision tree. The WEKA tool is used and the metrics used for classification are precision-recall, F measure and class, FP rate and TP rate.

SimranKanodia et al. [8] suggested a Markov Decision process for YouTube spam message detection and the result is compared to other data mining tools used in this field. The Markov Decision accuracy is 78.82 % which is quite better than those of other data mining algorithms out of which the maximum accuracy is obtained through the random forest which is 72.52 %.

Abdullah O. Abdullah et al [9] WEKA tool and python code is used for the employment of different classification algorithms on YouTube dataset that was generated using YouTube API. All the different 9 algorithms used have almost 90 % and above than that accuracy. Out of these different 9 algorithms, accuracy Adaptive Genetic Algorithm has performed quite well and achieved an accuracy of 99.1 %.

ShreyasAiyar et al. [10] in this they have used different machine learning algorithms along with custom approaches i.e. N-Grams. For automated detection of spam messages on YouTube videos they have suggested that the character gram approach performed better result as compared to word gram for obtaining better classification accuracy.

3. Outline of Spam Detection

S.No	Research Article	Techniques Used	Available Dataset	Results
1	N-Gram Assisted YouTube Spam Comment Detection	RF,SVM, Naïve Bayes,N-Gram	13000 Comments https://developers.google.com/youtube/v3/docs/commentThreads#Retrieve_comments [11]	N-Gram Outclassed Other Machine Learning Algorithms
2	A Comparative Analysis of Common YouTube Comment Spam Filtering Techniques	AGA, ICA-Amuse, ELM -AE, ANN-BP, SVM-K, K-NN, LR, NBC, DT	100 Channels having 10,000 Samples https://developers.google.com/youtube/v3/[12]	Adaptive Genetic Algorithm Performed better than other 8 Algorithms
3	A Novel Approach for YouTubeVideo Spam Detection using Markov Decision Process	Markov Decision Process, Decision Tree, Naïve Bayes, KNN, Random Forest, Ripper, Clustering	50 Videos, 2054 Instances out of which 824 was Spam Comment and 1230 normal	Markov Decision Process accuracy 78.52 % better than the best RF which was 72.82 %
4	Spam Detection Using KNN and Decision Tree Mechanism in Social Network	KNN, Decision Tree	FED Real Dataset http://mashable.com/2012/12/18/twitter-200-million-active-users/ Accessed July 22, 2016 [13]	Precision Call, F Measure and Class, FP and TP rate.
5	TubeSpam: Comment Spam Filtering on YouTube	Naïve Bayes, Decision Tree, Logistic Regression	5 Different YouTube Video Datasets Datasets / YouTube ID / # Spam/ # Ham / Total 1) Psy9bZkp7q19f0 175 175 350 2] KatyPerryCevxZvSJLk8 175 175 350 3) LMFAO KQ6zr6kCPj8 236 202 438 4) Eminemuel Hwf8o7 U 245 203 448 5) Shakira pRpeEdMmmQ0 174 196 370 http://dcomp.sor.ufscar.br/talmeida/youtubespamcollection/[14]	A confidence level of 99 % on all algorithms, Suggested their own App. Tube Spam.
6	A Data Mining Based Spam Detection System For YouTube	Naïve Bayes, Decision Tree, Clustering	Self Generated Data using Tube Kit http://www.tubekit.org/ [15] 1. No of Videos 1719 2. No of distinct users 1428 3. No of comments 10102865 4. No of ratings count 23013568 5. No of different categories 16	Lift Score is Calculated, Naive Bayes and Decision Tree performed better having an accuracy of 80.20 % and 82.11 % respectively

4 YouTube Data Generation Technique

This section discussed how YouTube comments data is generated from YouTube videos. The toolkit TubeKit [15] is used for customized YouTube Crawlers. This toolkit collects a variety of data from YouTube videos. The working design of the crawler is shown in the below fig 2.
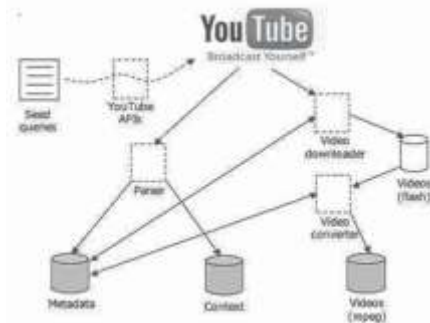

Fig.2. Scheme for Query Based YouTube Crawling [16]

5 Conclusion and Future

WorkYouTube a social networking feature website providing one of the largest video content publication. Recently a vast number of increases in his popularity and its new monetization policy for new channels have increased the number of users with low-quality information which is known as spam, which is posted solely for their channel promotion or to post the malicious video link.

The automatic moderation of comment is still an unexplored field and reason is the lack of automatic tool detection for spam filter messages. Due to which popular channelson YouTube try to avoid communication with their fans on this platform and use another platform for communication purpose.

The main goal behind this research was to find out the different techniques and strategies to find out the undesired comments i.e. spam messages and to describe different dataset available for the user working in this area. The results obtained through this research i.e. about techniques and available datasets description would further enhance the future comparison. Since for future work I would suggest that as there is not a single method that performs well on all the available dataset, so cascading of different machine learning classifiers along with best feature selection algorithms can be used to further enhance the result. Also preprocessing of the comment text can be done using natural language processing for text normalization. The reason behind text normalization is that the words used in comments are usually slang, idioms, emoticons, symbols, and abbreviations.

3 References

[1]   https://www.youtube.com/intl/en-GB/yt/about/press/

[2]   O'Callaghan, Derek et al. "Identifying Discriminating Network Motifs in YouTube Spam." CoRR abs/1202.5216 (2012): n. pag.

[3]   https://www.bbc.com/news/newsbeat-46559772

[4]   https://support.google.com/youtube/answer/111870?hl=en/

[5]   Chowdury, Rashid &Monsur Adnan, Md.Nuruddin& Mahmud, G.A.N. & Rahman, Mohammad. (2013). A data mining based spam detection system for YouTube. 373-378.

[6]   C. Alberto, Tulio&Lochter, Johannes & Almeida, Tiago. (2015). "TubeSpam: Comment Spam Filtering on YouTube." 138-143. 10.1109/ICMLA.2015.37.

[7]   Goyal, Saumya& K. Chauhan, R &Parveen, Shabnam.(2016)"Spam detection using KNN and decision tree mechanism in social network"522-526.10.1109/PDGC.2016.7913250.

[8]   S. Kanodia, R. Sasheendran and V. Pathari, "A Novel Approach for Youtube Video Spam Detection using Markov Decision Process," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 2018, pp. 60-66. doi: 10.1109/ICACCI.2018.8554405.

[9]   A. O. Abdullah, M. A. Ali, M. Karabatak and A. Sengur, "A comparative analysis of common YouTube comment spam

filtering techniques," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, 2018, pp. 1-5. doi: 10.1109/ISDFS.2018.8355315

[10] ShreyasAiyar, Nisha P Shetty, "N-Gram Assisted Youtube Spam Comment Detection", Procedia Computer Science, Volume 132,2018, Pages 174-182, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2018.05.181.

[11] Youtube, YoutubeDataAPI(v3), https://developers.google.com/youtube/v3/docs/commentThreads#Retrieve_comments

[12] Data API. Available at: https://developers.google.com/youtube/v3/ (Accessed on 22 November 2017)

[13] S. Fiegerman, "Twitter now has more than 200 million monthly active users," Mashable. [Online]. Available. http://mashable.com/2012/12/18/twitter-200-million-active-users/. Accessed July 22, 2016

[14] http://www.dt.fee.unicamp.br/~tiago//youtubespamcollection/

[15] TubeKit, http://www.tubekit.org/. , last accessed 24th August 2013.

[16] C. Shah, "Supporting Research Data Collection from YouTube with TubeKit". Proceedings of YouTube and 2008 Election Cycle in the United States, Amherst, MA: April 16-17, 2009.