Digital Forgeries: Problems and Challenges

Shrishail Math
Indian Institute of information Technology,
Allahabad, INDIA.

R.C.Tripathi
Indian Institute of Information Technology,
Allahabad, INDIA.

ABSTRACT

Today, we are leaving in digital era, over the past decade, digital Technology has matured to become predominant technology for creating, processing, transmitting and storing a information, a form of knowledge and intellectual assets. Multidimensional knowledge and intellectual assets are produced and represented in various forms such as audio, video, text , image , all together we can call it as a multimedia forms, finally all forms are stored as a digital bits and byte forms ie digital content .

The recent advances in software developments, plug and play(run) tools to capture, process, access and transmission of digitizes information, it has never so easy to alter the information without leaving any visual clues of tempering of digital data

The digital forgery is new research domain with many threats and opportunity with complexity in the problem; in this paper we discussed the seriousness of the problem, its impact and challenges ahead for future researchers

General Terms

Digital forgeries, Multimedia Security, Information Security and Assurance, Intellectual Property Protection.

Keywords

Digital forgeries, Multimedia forensics, digital forensics, digital information, data provenance, image forgery

1. INTRODUCTION

Today, we are living in digital era. Over the past decade, digital Technology has matured to become predominant technology for creating, processing, transmitting and storing a information, a form of knowledge and intellectual assets. Information a form of knowledge represented and created in multidimentional forms such as audio, video, text, image etc. digital technology has made it easy and possible to represent all and every form of information, that is knowledge and intellectual assets into a digitized form.

The digital technology undoubtedly superior to earlier conventional analog form and it has many advantages such as easy accessibility, searching, manipulation, transmission etc.

The recent advances in software developments by way of plug and play(run) tools to capture, process, access and transmission of digitizes information, has made it easy to alter the information without leaving any visual clues of tampering of digital data.

The forgeries are not new to human kind but very old problem. In past it was limited to art and literature but not affecting general public.

The digital technologies, multimedia, new set of digital acquisition,processing devices and tools and their easy availability, wide spread transmissions through social networking sites over web and open source software, all together has raised a serious and challenging problem of digital forgeries. It has raised new questions such as “ Is seeing believable?” “Are hearable voices authentic?” since these affecting the public in general and therefore there is urgent need to address the above questions

Fig 1 gives the glimpse of hierarchical overview of various forms of digital forgeries

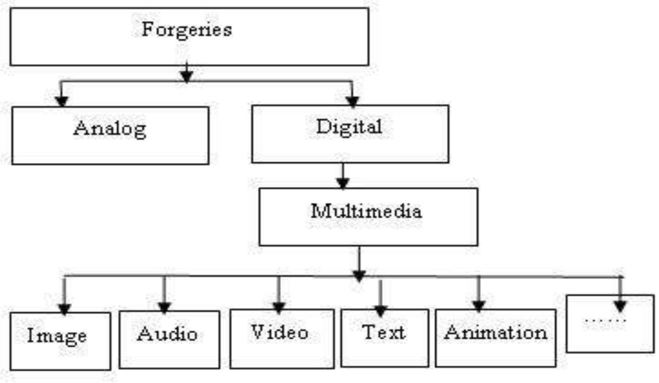

Fig 1: Hierarchical view of forgeries

2. RELATED WORK

The field of digital forgery had drawn the attention researcher worldwide. The problems being highlighted in this domain are digital forgeries of social impacts, detection techniques, and prevention techniques. The digital forgeries have many perspectives and implications on social, legal, technical, intelligence, investigative mechanisms, security, managerial issues.

Prof. Hany farid has lately drawn attention of researcher worldwide by reporting digital image forgery problem [1]. Later along with his students, he developed the techniques to detect the digital image forgery for specific cases [1, 2, 3].

3. FUNDAMENTALS

The forgery creation and detection are complimentary to each other. Fig2 explains the forgery creation and detection processes. The knowledge of forgery creation process contributes to the advances and sophistication in forgery detection methods and visa versa, both are never ending problems, the secrecy involved in both of the methods introduces a complexity in forgery creation as well forgery detection processes and acts as hindrance to both of these processes.

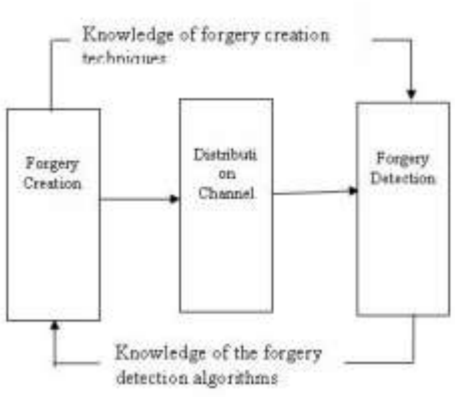
Knowledge of forgery creation techniques
Forgery Creation
Distributi on Channel
Forgery Detection
Knowledge of the forgery detection algorithms

Fig 2: Digital Forgery Creation

THE CURRENT RESEARCH PROBLEMS

The research in the field of digital forgeries is still in the infancy stage. Many innovative and promising technical problems of significant social importance are being identified. Also many realistic specific techniques, solutions have been proposed. The research in the domain of digital forensics is finally shaping towards the solving more generalized problem. Accordingly it emerging that generalized solutions and techniques, building standardized data sets, benchmarks, evaluation criteria etc are to be proposed to realize the new frameworks minimizing the chances for digital forgeries.

The fundamental problems in this fields are followings.

3.1 Natural or Computer generated? :

The images and data of financial, legal evidences, medical reports, such other a high valuable assets originality and authenticity is of prime importance. Identifying the originality and authenticity of image or data in many cases becomes challenging problem.

The advance in computer graphics, animation, multimedia in association of high computing machines, algorithms, increases the complexity of the issue. It is possible to generate high precision realistic images and data of any events.

Identifying and differentiating the data and image acquired by acquisition devices and realistic computer generated one is a multidimensional problem that has drawn attention of researchers worldwide.

3.2 Forgery detection

The easy availability of digital editing tools, alteration, and manipulation became very easy and as a result forgery detection becomes a complex and threatening problem. Fig 3 describes the basic operations involved in forgeries. Specific to image forgery detection image can be manipulated in various ways with many simple operations like affine transforms (such as translation, scaling, rotati*on*, shearing) compensation operations (like color, brightness, contrast adjustments, blurring and enhancement) suppression operation (such as filtering, compression and noise addition). Additionally more complex operations are also possible such as compositing, blending, matting, cropping, photomontage leading to visually untraceable artifacts in a image. The automatic and scientific method of detecting the forged images has become a biggest challenging problem to researchers and the same problem is true for every multimedia contents.

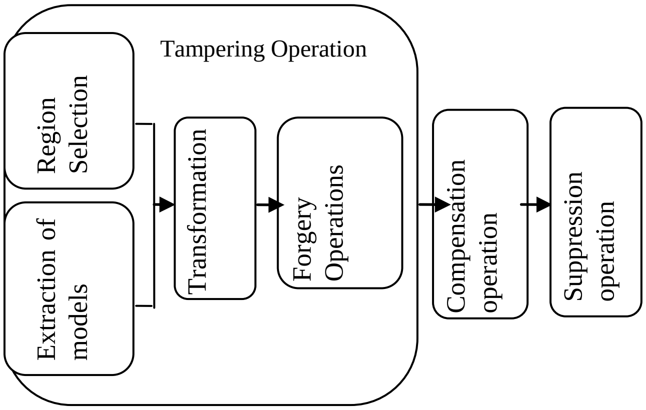

Fig 3: Operations in Digital Forgeries

3.3 Linage (Flow) Mapping

The high speed accessibility of internet and easy availability of freely available high processing digital editing tools (image) worsen the problem of authenticity of digital resources, the technology of digital resources is moving at a much faster rate due to social networking sites, its very difficulty find the origin of the resources. Thus finding the history (flow) of digital resources becomes a critical problem. Some efforts of finding the linage (flow) of data are being made in networked environment but as of our knowledge still nobody has tried to find the linage of digital resources. Researchers have not so for recognized as a potential problem in digital resources. In wake of finding solutions leads to solve today's leading problems related to the authenticity of the intellectual assets.

3.4 Source (Origin) Identification

In legal proceedings quite often, the proving authenticity of evidences and their sources becomes essential. The new varieties of image acquisition devices such as digital camera, scanners, cell phones, etc increase the complexity in the problem

4. CHALLENGES AHEAD

The biggest problem and challenge is how to ensure that intellectual assets in digital form are authentic and to tampered and their entire contents are authentic and consistent, the provenance of consistency, integrity authenticity (CIA) can only assure the digital intellectual assets origin and originality (unforged).

4.1 Data Provenance

The simmhan define data provenance as

"Information that helps determine the ***derivation history*** *of a data product, starting from its original sources"*

The data provenance is necessary for protection of rights and may be regulatory requirement in applications like science, medicine, financial transactions government legal prosecutions and many more daily situations, wherever the information is valuable and trustworthy.

Technological obsolescence: The technological developments in new devices, process, software and hardware is taking place at much faster rate than as assumed by Moore's Law. Most of existing digital technological infrastructures, methods and process of recording, storing, and retrieval of digital resources is being totally replaced in 2-5 years. That makes the preservation of digital evidences and investigations challenging issue.

The technological obsolescence represents a far greater threat to information in digital form than the inherent physical fragility of many digital media [5]

Formats: The intellectual assets in digital form available in various forms and each form having different file formats may … consistently changing technical evolution of hardware and software. The greater responsibility lies in preserving and achieving convertibility of one form to other formats and making the reservation formats independent of hardware and software and forward and backward interoperatability

4.2 Migration of digital information The preservation of integrity of digital documents over a migration of digital documents within organization and on internet with retaining capability of retrieval and display of integrated digital document poses greater challenge with constantly changing technology.

4.3 Ethical, legal and institutional issues

The wide spread uncertainties about legal, ethical and institutional issues of managing, preserving the intellectual assets (e.g., text and other document-like objects, photographs, film, software, multimedia objects) poses new threats and challenges.

4.4 Differentiating Forgery and clarity modification `

There are thinnest differences in identifying the forgeries and the alteration made in digital assets for the purpose of increasing the clarity of information representation. The biggest challenges are how to identify the few operations are made on digital assets to increase the visual clarity of documents without altering the meaning of digital assets and their origin.

4.5 Benchmarking and Standard data set

There is need of open data sets for critical and typical realistic conditions such as images (digital documents) in uncompressed with different resolutions, sizes and image acquisition model (camera model) with diverse contents for all possible forgeries such as copy paste, compositing, splicing, photomontage, blending, matting etc with manipulation, and manipulation compensation conditions like adjustments color, contrast, brightness, blurring, enhancement and possible post suppression operations like compression, recoloring addition of noise, etc. there is a need to evolve benchmarks for forged dataset as well as unforged data set in order to assess, evaluate, and understand the effectiveness of the research with collaborative studies.

4.6 Performance evaluation

The limited number of studies reported on forgery detection techniques and digital forgery detection research field is still in the state of infancy that might be reason for non development of performance evaluation bench marks, standardized data set, standard terminologies and techniques for determination of performances of existing techniques. There is need to borrow performance measurement criteria from other similar fields like fraud detection in credit cards, telecom frauds performance measurement terminology like true false rate, false positive, false negative, etc.

5. CONCLUSION AND FUTURE DIRECTIONS

In this paper, introduction and discussion is made of forgery research problems and associated challenges in solving a forgery problem. Increasingly significance of credibility of digital resources is becoming questionable? The need for assurance of authentic digital resources is more vital than ever. Ultimately it is hoped that due to increasing importance new significant problems of research and its associated challenges for researchers and practitioners in the digital communities soon become major area for research.

6. REFERENCES

[1].Hany Farid" .2009, Image Forgery Detection A Survey", IEEE Signalprocessing Magazine

[2] M. K. Johnson and H. Farid,2005, "Exposing digital forgeries by detecting inconsistencies in lighting," in Proc. ACM Multimedia and Security Workshop, NewYork, NY, pp. 1–10.

[3] Lukas J, Fridrich J, Goljan M.,” 2005,Determining digital image origin using sensor imperfections”, In: Proceedings of SPIE Electronic Imaging,Image and Video Communications and Processing, 5685(1): 249–260.

[4] S. Ye, Q. Sun, and E. C. Chang,”2007 Detecting digital image forgeries by measuring inconsistencies of blocking artifact,”, in Proc. IEEE Int. Conf. Multimedia and Expo, Beijing, China, pp. 12–15.

[5] Mallinson, John C. ,1986, “Preserving Machine-Readable Archival Records for the Millenia.” Archivaria 22(Summer): 147-52.

[6]S.Bayram, I. Avcibas, B. Sankur, and N. Memon,,” 2005,“Image manipulation detection with binary similarity measures,” in Proc. European Signal Processing Conf., Turkey,